INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2004/011306

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl[7] C12N15/56, C12N5/10, C12P19/04, A01H5/00, C08B37/08, A61K31/728

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl[7] C12N15/56, C12N5/10, C12P19/04, A01H5/00, C08B37/08, A61K31/728

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
JSTPlus(JOIS), SwissProt/PIR/GeneSeq, Genbank/EMBL/DDBJ/GeneSeq, BIOSIS/WPI(DIALOG)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X Y	Graves M.V. et al., Hyaluronan synthesis in virus PBCV-1-infected chlorella-like green algae, Virology, 1999, Vol.257, No.1, pages 15 to 23	20 1-19
X Y	DeAngelis P.L. et al., Hyaluronan synthase of chlorella virus PBCV-1, Science, 1997, Vol.278, No.5344, pages 1800-3	20 1-19
Y	JP 2001-521741 A (THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA), 13 November, 2001 (13.11.01)	1-19

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
29 September, 2004 (29.09.04)	12 October, 2004 (12.10.04)
Name and mailing address of the ISA/ Japanese Patent Office Facsimile No.	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (January 2004)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2004/011306

Box No. I Nucleotide and/or amino acid sequence(s) (Continuation of item1.b of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application and necessary to the claimed invention, the international search was carried out on the basis of:

   a. type of material

   [x] a sequence listing

   [ ] table(s) related to the sequence listing

   b. format of material

   [ ] in written format

   [x] in computer readable form

   c. time of filing/furnishing

   [ ] contained in the international application as filed

   [x] filed together with the international application in computer readable form

   [ ] furnished subsequently to this Authority for the purposes of search

2. [x] In addition, in the case that more than one version or copy of a sequence listing and/or table relating thereto has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that in the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2004/011306

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:
(See extra sheet.)

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☒ No protest accompanied the payment of additional search fees.

Claims 1, 3, 6 to 8, 14, 16 and 18 relate to a method of producing hyaluronic acid by transforming a plant cell with the use of a vector containing a DNA encoding hyaluronate synthase and then growing the plant cell; claims 2, 4 to 5, 9 to 13, 15, 17 and 19 relate to a method of producing hyaluronic acid by transforming a plant with the use of a vector containing a DNA encoding hyaluronate synthase and then growing the plant; and claim 20 relates to hyaluronic acid.

As the results of the search, however, it is found out that "a method of producing hyaluronic acid by transforming a host with the use of a vector containing a DNA encoding hyaluronate synthase and then growing the host" is not novel because of having been reported in document 'JP 2001-521741 A (THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA) 13 November, 2001 (13.11.01), Claims 39 to 41'.

As a result, "a method of producing hyaluronic acid by transforming a host with the use of a vector containing a DNA encoding hyaluronate synthase and then growing the host" falls within the category of prior art and, therefore, the above common matter cannot be referred to as a special technical feature in the meaning within the second sentence of PCT Rule 13.2.

Thus, there is no matter common to all claims.

Since there is no other common matter seemingly being a special technical feature in the meaning within the second sentence of PCT Rule 13.2, no technical relevancy in the meaning within PCT Rule 13 can be found out among these groups of inventions differing from each other.

Such being the case, it is obvious that claims 1 to 20 do not comply with the requirement of unity of invention.

Virology 257, 15–23 (1999)
Article ID viro.1999.9628, available online at http://www.idealibrary.com on IDEAL®

Hyaluronan Synthesis in Virus PBCV-1-Infected Chlorella-like Green Algae[1]

Michael V. Graves,* Dwight E. Burbank,* Robyn Roth,† John Heuser,† Paul L. DeAngelis,‡ and James L. Van Etten*[2]

*Department of Plant Pathology, University of Nebraska, Lincoln, Nebraska 68583-0722; †Department of Cell Biology and Physiology, Washington University Medical School, St. Louis, Missouri 63130; and ‡Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma 73190

Received January 14, 1999; returned to author for revision January 20, 1999; accepted February 1, 1999

We previously reported that the chlorella virus PBCV-1 genome encodes an authentic, membrane-associated glycosyltransferase, hyaluronan synthase (HAS). Hyaluronan, a linear polysaccharide chain composed of alternating β1,4-glucuronic acid and β1,3-*N*-acetylglucosamine groups, is present in vertebrates as well as a few pathogenic bacteria. Studies of infected cells show that the transcription of the PBCV-1 *has* gene begins within 10 min of virus infection and ends at 60–90 min postinfection. The hyaluronan polysaccharide begins to accumulate as hyaluronan-lyase sensitive, hair-like fibers on the outside of the chlorella cell wall by 15–30 min postinfection; by 240 min postinfection, the infected cells are coated with a dense fibrous network. This hyaluronan slightly reduces attachment of a second chlorella virus to the infected algae. An analysis of 41 additional chlorella viruses indicates that many, but not all, produce hyaluronan during infection. © 1999 Academic Press

Key Words: glycosyltransferase; hyaluronan synthase; dsDNA virus; PBCV-1; chlorella virus; Phycodnaviridae.

INTRODUCTION

Hyaluronan, also called hyaluronic acid, is the least complex member of the glycosaminoglycan family, which also includes heparin, heparan sulfate, chondroitin, and keratan sulfate. The latter members of this family are highly sulfated and typically exist as a proteoglycan (i.e., covalently attached to a core protein). Hyaluronan, however, is a simple linear polysaccharide chain composed of alternating β1,4-glucuronic acid (GlcA) and β1,3-*N*-acetylglucosamine (GlcNAc) moieties that can reach molecular masses of up to 10^7 kDa (~25,000 disaccharides) (Hascall *et al.*, 1994; Laurent and Fraser, 1992). Hyaluronan is a ubiquitous constituent of the extracellular matrix, particularly of soft connective tissues in vertebrates (Laurent and Fraser, 1992). Hyaluronan interacts with proteins such as CD44 (Aruffo *et al.*, 1990; Culty *et al.*, 1990; Miyaka *et al.*, 1990), RHAMM (Hall *et al.*, 1995; Hardwick *et al.*, 1992), and BEHAB (Jaworski *et al.*, 1994). Consequently, this polysaccharide influences the growth and migration of cells in such diverse processes as embryonic development (Toole, 1991), oocyte maturation (Salustri *et al.*, 1990), angiogenesis, wound healing (West *et al.*, 1985), and tumor progression (Sherman *et al.*, 1994). In contrast to other glycosaminoglycans, which are assembled as they traverse the endoplasmic reticulum and the Golgi stacks, hyaluronan is synthesized by an enzyme located on the inner surface of the plasma membrane (Philipson and Schwartz, 1984). Hyaluronan synthase (HAS) adds sugar residues from UDP-GlcA and UDP-GlcNAc. In animal cells, hyaluronan is transferred to the pericellular space.

Extracellular capsules of a few pathogenic bacteria such as group A and C *Streptococcus* spp. and *Pasteurella multocida* also contain hyaluronan (Carter and Annau, 1953; Kass and Seastone, 1944). Because hyaluronan, a host component, is not normally immunogenic, the capsule serves as a molecular camouflage protecting the microbes from phagocytosis and complement fixation during infection (Husmann *et al.*, 1997; Schmidt *et al.*, 1996).

While sequencing the 330,740-bp genome of the algal virus PBCV-1 (Kutish *et al.*, 1996; Li *et al.*, 1995, 1997; Lu *et al.*, 1995, 1996), we discovered that this virus contains an open reading frame (ORF) (A98R) that encodes a protein with similarity to both vertebrate and bacterial HAS enzymes. The PBCV-1 *has* gene was expressed in *Escherichia coli*, and the recombinant protein was an authentic, membrane-associated HAS (DeAngelis *et al.*, 1997). Landstein *et al.* (1998) demonstrated that PBCV-1 encoded two other enzymes, glutamine:fructose-6-phosphate amidotransferase (GFAT, ORF A100R) and UDP-glucose dehydrogenase (UDP-GlcDH, ORF A609L), that produce sugar precursors (glucosamine-6-phosphate and UDP-glucuronic acid, respectively) required for hyaluronan synthesis. In the current work, we monitor the expres-

[1] This manuscript has been assigned Journal Series No. 12512; Agricultural Research Division, University of Nebraska.

[2] To whom reprint requests should be addressed. Fax: (402) 472-2853. E-mail: vanetten@biocomp.unl.edu.

0042-6822/99 $30.00

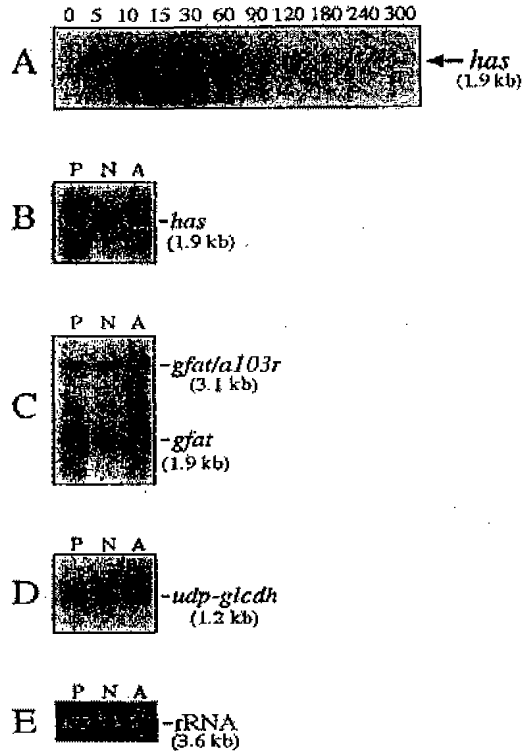


FIG. 1. Northern blot analysis of the accumulation of *has*, *gfat*, and *udp-glcdh* mRNAs during virus PBCV-1 infection. (A) RNAs isolated from uninfected (lane 0) and from PBCV-1-infected chlorella cells at 5, 10, 15, 30, 60, 90, 120, 180, 240, and 300 min p.i. The blot was probed with the PBCV-1 *has* gene. (B–D) RNAs isolated from cells at 30 min after infection with viruses PBCV-1, NC-1C, or AL-2C (lanes P, N, and A, respectively). The membranes were probed with either the PBCV-1 *has* gene (B), *gfat* gene (*a100r*) (C), or *udp-glcdh* gene (*a609l*) (D). The filter in panel E was stained with ethidium bromide and shows the 3.6-kb rRNA used to correct for loading differences between the samples.

sion of the *has* gene and the production and localization of hyaluronan in PBCV-1-infected chlorella.

RESULTS

The *has* gene expression during PBCV-1 replication

Total RNA was isolated from chlorella cells at various times after PBCV-1 infection and hybridized to the viral *has* gene. A single, ~1900-nucleotide transcript, a size sufficient to encode a protein of 567 amino acids, appeared within 5–10 min postinfection (p.i.), peaked at 30 min p.i., and disappeared at 60–90 min p.i. (Fig. 1A). Because PBCV-1 DNA synthesis begins ~60 min p.i. (Van Etten *et al.*, 1984), the *has* gene is an early gene. This is consistent with our previous finding that HAS enzyme activity was detected in chlorella cells at 50 and 90 min p.i. (DeAngelis *et al.*, 1997).

Hyaluronan is localized on the surface of PBCV-1-infected chlorella

Typically, HASs are integral membrane-bound proteins, and the newly synthesized hyaluronan is secreted across the membrane to the extracellular matrix (Philipson and Schwartz, 1984). Previous experiments established that the PBCV-1-encoded HAS is associated with the membrane fraction of PBCV-1-infected chlorella cells (DeAngelis *et al.*, 1997). Therefore, we looked for hyaluronan on the surface of infected chlorella by monitoring the ability of a ^{125}I-labeled hyaluronan-binding protein (^{125}I-HABP) to interact with intact, virus-infected cells (Tengblad, 1980). This protein did not attach to uninfected cells, indicating that the cell surface lacked hyaluronan. By 15 min p.i., small but significant amounts of the ^{125}I-HABP bound to the infected cells, indicating the presence of surface hyaluronan. During the first 90 min p.i., the level of ^{125}I-HABP bound to the infected cells increased slightly and then increased rapidly during the next 120–150 min (Fig. 2). Treatment of infected chlorella cells at 240 min p.i. with hyaluronan-lyase, before the addition of ^{125}I-HAPB, reduced attachment of the binding protein to the level of infected cells at 15 min p.i. (Fig. 2). The absolute specificity of the HABP and the hyaluronan-lyase for hyaluronan establish the presence of hyaluronan on the infected cell surface.

To determine whether the hyaluronan is localized to a specific area of the cell wall or is present over the entire cell surface, hyaluronan accumulation was also monitored by fluorescent microscopy using biotinylated-hyaluronan binding protein (bt-HABP) in conjunction with an avidin–FITC conjugate. As shown in Fig. 3A, many infected cells developed a uniform green fluorescence over the entire cell surface by 30 min p.i.; uninfected cells autofluoresced orange-red. The intensity of green fluorescence as well as the number of fluorescing cells increased up to 240 min p.i. Treatment of cells at 240 min p.i. with hyaluronan-lyase, before the addition of bt-HABP, abolished most of the green fluorescence (Fig. 3A).

Ultrastructural changes in the cell wall of PBCV-1-infected chlorella cells

The cell walls of uninfected and PBCV-1-infected cells were also examined by quick-freeze deep-etch electron microscopy. As shown in Fig. 3B, the exterior surface of the infected chlorella cell wall takes on a "hairy" appearance; by 240 min p.i., the infected cell is covered with a highly developed, dense fibrous network. Incubation of cells with hyaluronan-lyase removes this "hairy" material, indicating that this fibrous network is composed of hyaluronan (Fig. 3B).

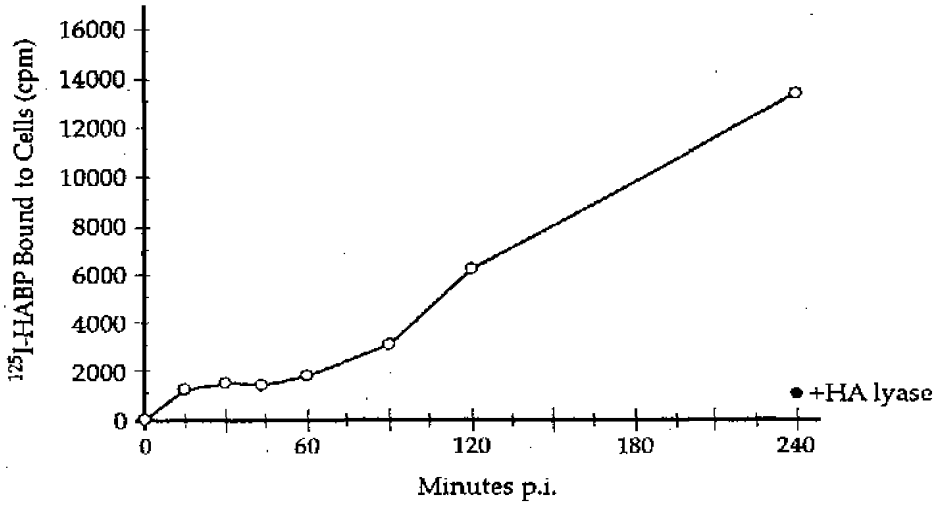


FIG. 2. Hyaluronan accumulation on the surface of PBCV-1 infected algae. Uninfected and infected cells were incubated with ^{125}I-HABP as described in Materials and Methods. One sample from infected cells at 240 min p.i. was treated with hyaluronan-lyase before the addition of the HABP. The results represent the average of two separate experiments.

Hyaluronan partially blocks chlorella virus attachment

The results of the preceding experiments indicate that the external surface of the chlorella cell wall changes dramatically after PBCV-1 infection as the result of hyaluronan accumulation. To determine whether surface hyaluronan prevents attachment of a second virus to PBCV-1-infected cells, we monitored the ability of an antigenic variant of PBCV-1 [named P31 (Wang *et al.*, 1993)] to attach to chlorella cells at various times after PBCV-1 infection. As shown in Table 1, prior PBCV-1 infection of chlorella reduces P31 attachment to the alga by ~50% at 120–300 min p.i. Treatment of the 180-min PBCV-1-infected cells with hyaluronan-lyase before the addition of P31 slightly increased the ability of P31 to attach to the cells (Table 1). Therefore, surface hyaluronan accumulation slightly reduces subsequent virus attachment, albeit late in the infection cycle.

The *has* gene is widespread in the chlorella viruses

To determine whether the *has* gene is widespread among the chlorella viruses, the *has* gene probe used in Fig. 1 was hybridized to DNA from 41 other viruses isolated from diverse geographical regions (Fig. 4). These viruses infect either *Chlorella* NC64A or *Chlorella* Pbi. Chlorella cells infected with each of the viruses were also monitored for extracellular hyaluronan with the ^{125}I-HABP (radioactive counts are also listed in Fig. 4). These experiments produced the following results. (1) The PBCV-1 *has* gene probe did not hybridize to host *Chlorella* NC64A DNA. (2) The PBCV-1 *has* gene probe hybridized to some degree to 28 of the 37 DNAs from viruses (including PBCV-1) that infect *Chlorella* NC64A (NC64A viruses). (3). All except one of these 28 NC64A viruses produced extracellular hyaluronan. The exception was NY-2A, which hybridized weakly with the *has* gene. (4) Nine of the 37 NC64A viruses, CA-1A, CA-2A, IL-2A, IL-2B, IL-3A, IL-3D, SC-1A, SC-1B, and IL-5–2s1, neither hybridized with the *has* gene probe nor produced extracellular hyaluronan. (5) None of the DNAs from the five viruses, CVA-1, CVB-1, CVG-1, CVM-1, and CVR-1, that infect *Chlorella* strain Pbi (Pbi viruses) (Reisser *et al.*, 1988) hybridized with the PBCV-1 *has* gene probe. However, 2 of the 5 Pbi viruses, CVG-1 (Fig. 3C) and CVR-1, produced extracellular hyaluronan (Fig. 4). Presumably, these 2 Pbi viruses encode a *has* gene that has diverged substantially from the PBCV-1 gene.

These experiments indicate that *has* gene expression is not essential for chlorella virus growth because 10 of the NC64A viruses and 3 of the Pbi viruses do not produce detectable extracellular hyaluronan. One explanation is that these viruses encode an enzyme or enzymes that produce another polysaccharide on the external surface of the infected chlorella cells. However, two experiments indicate that this possibility is unlikely. (1) The surface of chlorella cells infected with Pbi virus CVA-1 (no ^{125}I-HABP binding; Fig. 4) does not appear "hairy" on electron microscopy at 240 min p.i. (Fig. 3C). (2) Cells infected with virus IL-3A were also monitored for changes in attachment of a second virus; only a slight reduction in attachment occurred (results not shown).

Analysis of *has* gene sequence and expression from other chlorella viruses

Ten of the 41 chlorella viruses, plus PBCV-1, were chosen for further analysis of the *has* gene. The *has* gene from each was amplified by PCR using primers that

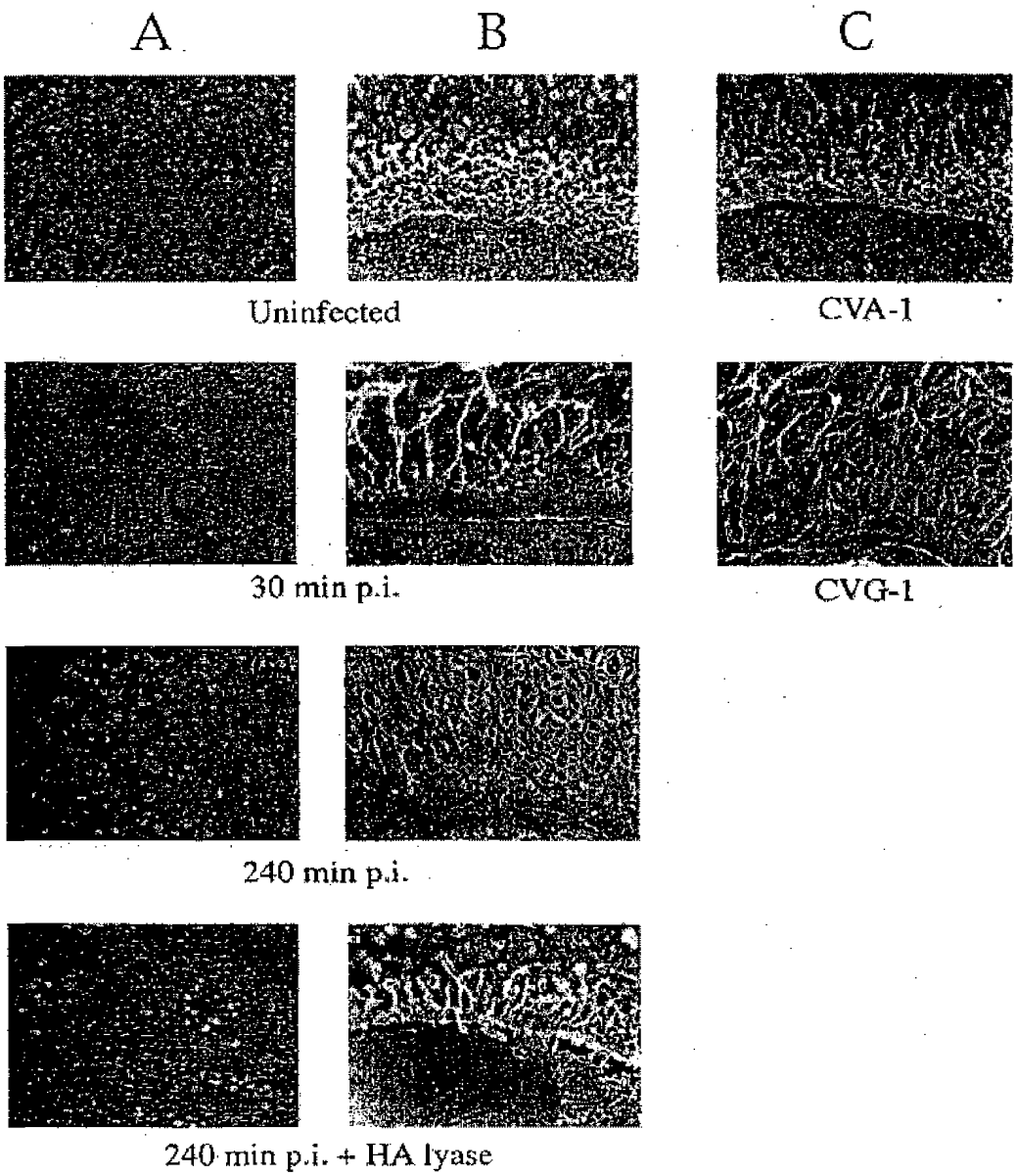


FIG. 3. Localization of hyaluronan on the surface of infected cells and ultrastructural changes in the algal cell wall after viral infection. Detection of hyaluronan on the surface of uninfected *Chlorella* NC64A cells and PBCV-1-infected cells at 30 and 240 min p.i. using a bt-HABP in combination with an avidin–FITC conjugate (A) or by quick-freeze deep-etch electron microscopy (B). A sample of PBCV-1-infected cells at 240 min p.i. was treated with hyaluronan-lyase before the fluorescent labeling or electron microscopic examination. Note the bright green fluorescence and accumulation of a dense, fibrous network on the surface of PBCV-1-infected cells at 240 min p.i. (C) *Chlorella* Pbi cells infected with viruses CVA-1 or CVG-1 examined by quick-freeze, deep-etch electron microscopy. Both samples were taken at 240 min p.i. Note the surface of the CVA-1-infected cells resemble uninfected *Chlorella* NC64A cells as contrasted to the surface of CVG-1- and PBCV-1-infected cells.

correspond to the 5′ and 3′ ends of the PBCV-1 *has* gene (DeAngelis *et al.*, 1997). The primers produced the expected 1.7-kb product from virus PBCV-1 DNA and five additional virus DNAs, NC-1C, AL-2C, MA-1E, CA-4A, and XZ-5C (Fig. 5). No PCR product was obtained with SC-1A, MA-1D, NY-2B, NY-2A, and CVG-1 DNAs. These results support the data in Fig. 4; PCR products were produced only from virus DNAs that hybridized strongly to the PBCV-1 *has* gene probe. The six different 1.7-kb PCR products (including PBCV-1) were cloned and sequenced; analyses of the sequences led to the following conclusions. (1) All the clones (including PBCV-1) contained a G instead of an A at position 52285 in the original PBCV-1 genomic sequence (Li *et al.*, 1995), indicating an error in the published PBCV-1 sequence. Correction of this base changes an Asp residue to a Gly at

TABLE 1

Attachment of Virus P31 (an Antigenic Variant of PBCV-1) to PBCV-1-Infected *Chlorella* NC64A

Time after PBCV-1 infection (min)	Percentage of unattached P31*
15	19 ± 8
60	28 ± 8
120	36 ± 7
180	40 ± 3
240	34 ± 8
300	32 ± 8
180 + HA-lyase	29 ± 10

* Average of three separate experiments.

amino acid 462. This change is significant because all other eukaryotic HASs have a Gly in this position (DeAngelis *et al.*, 1997). (2) The sequence of the AL-2C clone was identical to the corrected PBCV-1 sequence. (3) The

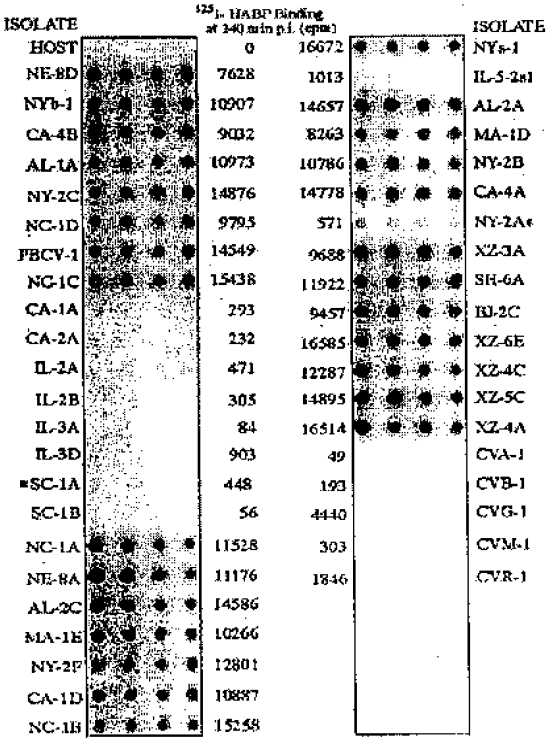
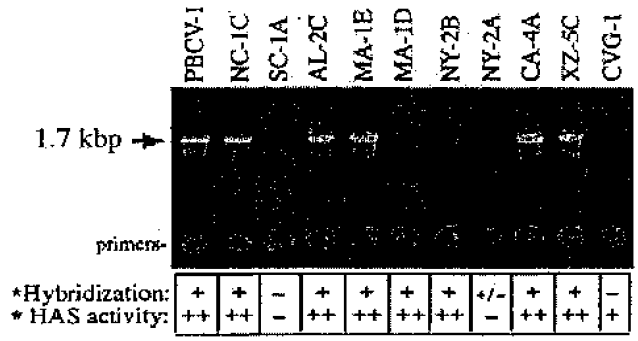


FIG. 4. Hybridization of the PBCV-1 *has* gene to DNA isolated from *Chlorella* NC64A and from 37 NC64A viruses and 5 Pbi viruses (CVA-1, CVB-1, CVG-1, CVM-1, and CVR-1). The blots contain 1, 0.5, 0.25, and 0.12 μg of DNA (left to right, respectively). The accumulation of hyaluronan on the surface of the infected cells, as measured by the ability of ^{125}I-HABP to attach to the cells at 240 min p.i., is also indicated for each virus. The labeling results represent the average of at least two separate experiments. Because viruses SC-1A and NY-2A replicate slower than the other viruses, they were analyzed for hyaluronan accumulation at 8 h p.i. (*).



* Data summarized from Figure 4

FIG. 5. PCR products produced from PBCV-1 and 10 additional chlorella viruses using primers derived from the PBCV-1 *has* gene sequence. (Bottom) Dot-blot hybridization and ^{125}I-HABP binding data from Fig. 4. No PCR products were obtained from virus DNAs that either failed to hybridize (SC-1A and CVG-1) or hybridized poorly (NY-2A, NY-2B, and MA-1D) to the PBCV-1 *has* gene probe.

sequences of the MA-1E and XZ-5C clones each contained a single, distinct, silent nucleotide difference from PBCV-1. (4) The sequence of the NC-1C clone varied by 11 nucleotides from PBCV-1; four of these changes resulted in amino acid substitutions. Three of these substitutions (R136→K, D400→E, and V534→I) were conservative changes, whereas the fourth (T360→A) was nonconservative. (5) The sequence of the CA-4A clone differed from PBCV-1 by 15 nucleotides; eight of these differences were in common with NC-1C. Seven of these differences (two were in the same codon) led to six amino acid changes; four (R136→K, I450→V, V534→I, and T562→S) were conservative, whereas two (V529→T and E554→K) were nonconservative.

Total RNA was isolated from cells infected with NC-1C or AL-2C viruses at 30 min p.i. (the time when the *a98r* gene transcript is most abundant in PBCV-1-infected cells) and analyzed by Northern blotting with the PBCV-1 *has* gene probe (Fig. 1B) as well as probes for the two PBCV-1 genes (*gfat* and *udp-glcdh*) that encode enzymes synthesizing hyaluronan precursors (Figs. 1C and 1D). This experiment led to the following results. (1) Like PBCV-1, the *has* gene probe hybridized to a 1.9-kb RNA from cells infected with each virus. (2) The *gfat* and *udp-glcdh* probes produced the same hybridization patterns for all three viruses. Landstein *et al.* (1998) demonstrated that in PBCV-1, the largest of the three RNAs detected by the *gfat* probe results from readthrough transcription of the *gfat* gene into the adjacent *a103r* gene. The *a103r* gene encodes an mRNA capping enzyme (Ho *et al.*, 1996). (3) Although the amounts of *has*, *gfat*, and *udp-glcdh* mRNAs that accumulated in AL-2C-infected cells was approximately equal to that in PBCV-1-infected cells, the amount of *has*, *gfat*, and *udp-glcdh* mRNAs that accumulated in NC-1C-infected cells was reduced considerably. Therefore, all three viral genes involved in hyaluronan synthesis are expressed in two

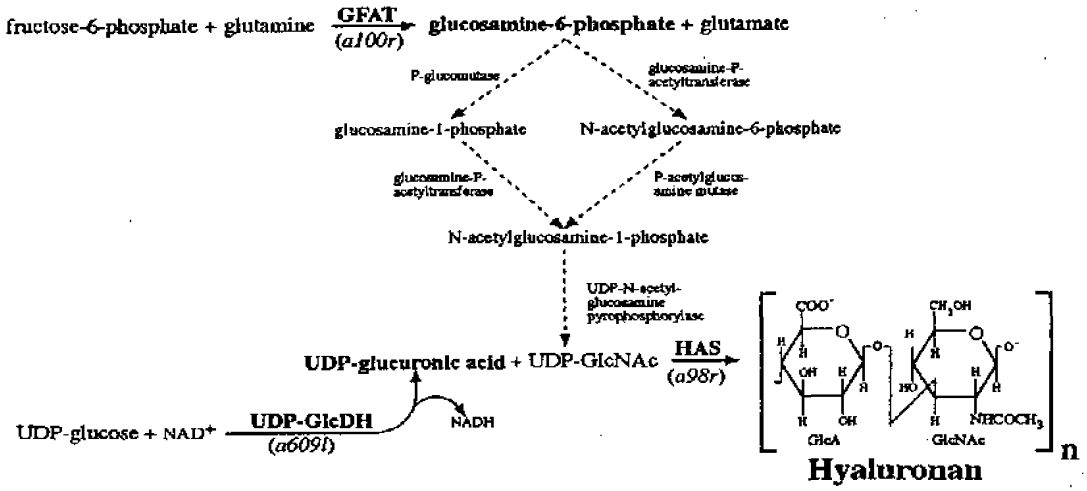


FIG. 6. Biosynthesis of hyaluronan starting with fructose-6-phosphate and UDP-glucose. Virus PBCV-1 encodes the enzyme GFAT (*a100r*), UDP-GlcDH (*a609l*), and HAS (*a98r*). The conversion of glucosamine-6-phosphate to UDP-GlcNAc requires at least three additional steps, designated by the dashed arrows. The genetic sources of these three additional enzymes are unknown.

other chlorella viruses, albeit at variable levels. This variation could reflect slight differences in the length of the NC-1C infection cycle compared with PBCV-1 and AL-2C.

DISCUSSION

We previously reported that chlorella virus PBCV-1 encodes an authentic, membrane-associated HAS (DeAngelis *et al.*, 1997). The PBCV-1 *has* gene was expressed in *E. coli*, and as expected, the recombinant protein required the simultaneous presence of UDP-GlcA and UDP-GlcNAc and manganese for activity. Hyaluronan lyase degrades the 3–6 × 10^6-kDa hyaluronan polysaccharide product of the recombinant enzyme. Results presented here establish that the PBCV-1 *has* gene is expressed within 10 min after PBCV-1 infection and that large amounts of hyaluronan accumulate on the cell surface of infected algae. To our knowledge, PBCV-1 is the first virus to encode an enzyme that synthesizes a polysaccharide. Viruses generally use host-encoded glycosyltransferases to create new glycoconjugates or accumulate host cell glycoconjugates during virion maturation. An ecdysteroid UDP-glucosyltransferase encoded by several baculoviruses is the only previously known virus-encoded glycosyltransferase with a characterized activity (O'Reilly, 1995). The ecdysteroid UDP-glucosyltransferase inactivates the insect's molting hormones by adding a single glucose residue to the hormone.

PBCV-1 probably encodes several other glycosyltransferases in addition to HAS. Studies on four PBCV-1 antigenic variants with altered oligosaccharide moieties on the three virion-associated glycoproteins led to the prediction that PBCV-1 encodes at least part, if not its entire, glycosylation machinery (Que *et al.*, 1994; Wang *et al.*, 1993). However, several observations indicate that the HAS enzyme described in this report does not glycosylate the PBCV-1 glycoproteins. (1) The oligosaccharide or oligosaccharides attached to the PBCV-1 glycoproteins contains only neutral sugars, glucose, galactose, mannose, fucose, xylose, rhamnose, and arabinose (Wang *et al.*, 1993). (2) Hyaluronan accumulates on the outside of the virus-infected host, whereas intact infectious virus particles accumulate inside the host at least 30–40 min before release by lysis of the host cell wall. (3) Typically, hyaluronan is not covalently bound to a protein (Hascall *et al.*, 1994; Laurent and Fraser, 1992). Therefore, we conclude that HAS is not involved in PBCV-1 protein glycosylation and that the virus encodes separate glycosyltransferases for this purpose.

Landstein *et al.* (1998) previously demonstrated that PBCV-1 encodes two additional enzymes involved in hyaluronan biosynthesis: GFAT and UDP-GlcDH (Fig. 6). UDP-GlcDH converts UDP-glucose into UDP-GlcA, a precursor of hyaluronan. GFAT converts fructose-6-phosphate into glucosamine-6-phosphate, an intermediate in UDP-GlcNAc biosynthesis. Like *has*, the *udp-glcdh* and *gfat* genes are expressed early in PBCV-1 infection (Landstein *et al.*, 1998). At least three additional enzymes are needed to convert glucosamine-6-phosphate into UDP-GlcNAc (Fig. 6). Sequence comparisons have failed to identify candidate genes encoding these enzymes in the viral genome. However, the presence of three virus-encoded hyaluronan biosynthetic enzymes suggests that the polysaccharide serves an important function in the PBCV-1 life cycle.

The extracellular hyaluronan does not play an obvious role in the interaction between PBCV-1 and its algal host because neither plaque size nor plaque number is al-

tered by including either testicular hyaluronidase or free hyaluronan in the top agar of the PBCV-1 plaque assay (DeAngelis *et al.*, 1997). However, the extracellular hyaluronan weakly inhibits attachment of additional viruses, especially late in infection (Table 1). This inhibition might be advantageous to PBCV-1 because it would reduce multiple infections. The significance of this hyaluronan-mediated reduction in virus attachment is questionable, however, because NC64A viruses mutually exclude one another by a hyaluronan-independent mechanism (Chase *et al.*, 1989). This exclusion phenomenon, which has no effect on virus attachment, occurs before hyaluronan-mediated inhibition of virus attachment.

We considered two other biological functions for the PBCV-1-encoded hyaluronan; these functions are based on our limited knowledge of the natural history of the viruses. Chlorella viruses are ubiquitous in freshwater collected worldwide, and titers as high as 4×10^4 infectious viruses/ml of native water have been reported (Van Etten *et al.*, 1985; Yamada *et al.*, 1991). The only known hosts for these viruses are chlorella-like green algae that normally live as hereditary endosymbionts in some isolates of the ciliate *Paramecium bursaria*. In the symbiotic unit, algae are enclosed individually in perialgal vacuoles and are surrounded by a host-derived membrane (Reisser, 1992). The initial establishment and the long-term maintenance of symbiosis require that the algae avoid digestion by the paramecium. Reassociation studies with different *Chlorella* spp. and algae-free *P. bursaria* indicate that only the original symbiotic algae effectively reestablish a long-term, stable symbiosis with the ciliate (Reisser, 1992). Other chlorella species are digested. Presumably, the relationship between the algae and the paramecia require interactions of specific algal surface components with host membrane factors (Meints and Pardy, 1980; Pool, 1979). Interestingly, endosymbiotic chlorella are resistant to virus infection during symbiosis and become infected only when they are grown outside the paramecia (Reisser *et al.*, 1991).

One possible biological function for hyaluronan is that polysaccharide accumulation on the algal surface inhibits the uptake of virus-infected algae by paramecium. Prevention of the internalization of infected algae would enhance virus survival because virions released inside the paramecium would presumably be destroyed by the protozoan's digestive system. Alternatively, the chlorella viruses might have another host in nature; perhaps the virus is transmitted because this other host is attracted to or binds to hyaluronan on virus-infected algae. In this regard, it is interesting that the intestinal pathogen *Entamoeba histolytica* has a surface protein that binds to hyaluronan (Renesto *et al.*, 1997).

However, complicating the issue of the biological significance of the extracellular hyaluronan in the PBCV-1 life cycle is the finding that some chlorella viruses lack the *has* gene and do not produce extracellular hyaluronan. Furthermore, cell walls of the chlorella host infected with these viruses do not take on a "hairy" appearance. Consequently, the extracellular production of hyaluronan or an equivalent extracellular polysaccharide is not essential for survival of the viruses in nature because all of the tested chlorella viruses have been isolated from natural sources within the past 18 years. In contrast, all the *Chlorella* NC64A viruses encode the *gfat* and *udp-glcdh* genes, as judged by dot-blot analysis (Landstein *et al.*, 1998).

MATERIALS AND METHODS

Chlorella, viruses, and plasmids

The hosts for the chlorella viruses, *Chlorella* strain NC64A and *Chlorella* strain Pbi, were grown on MBBM medium (Van Etten *et al.*, 1983) and FES medium (Reisser *et al.*, 1988), respectively. Procedures for producing, purifying, and plaquing virus PBCV-1 and the other chlorella viruses and isolating host and virus DNAs have been described (Van Etten *et al.*, 1981, 1983, 1983a). The plasmid pCVHAS, which contains the PBCV-1 *has* gene, has also been described (DeAngelis *et al.*, 1997).

Detection of hyaluronan on the surface of infected cells

Virus-infected cells used to measure hyaluronan accumulation were obtained by concentrating 1.5×10^7 cells/ml to 2.0×10^9 cells/ml, infection with PBCV-1 (m.o.i. of 5), and collection of 2.0×10^9 cells at various times p.i. Hyaluronan was detected on the surface of intact, infected cells using ^{125}I-HABP (Pharmacia Biotech, Uppsala, Sweden). Approximately 0.1 μCi of ^{125}I-HABP was added to the infected cells, which were then incubated on ice for 60 min. The cells were collected by centrifugation, and the supernatant containing unbound, labeled protein was removed. The amount of radioactivity (i.e., the amount of HABP bound to the cells) was determined with a gamma counter. Fluorescent visualization of hyaluronan on the surface of intact cells was accomplished by adding 1.5 μg of biotinylated aggrecan, a hyaluronan-specific binding protein (Applied Bioligands Co., Winnipeg, Canada) to 2.0×10^6 cells in 100 μl and incubating on ice 60 min. The cells were washed three times in PBS, resuspended in 100 μl avidin-FITC conjugate diluted 1:2000 in PBS (Sigma Chemical Co., St. Louis, MO) followed by an additional 60-min incubation on ice. The cells were then washed three times in PBS, resuspended in 20–50 μl of PBS, and examined under UV illumination with a Zeiss Axioskop UV microscope. In some experiments, duplicate samples were treated with 10–50 units of hyaluronan-lyase (Sigma Chemical Co.) for 60 min before the addition of the HABP. Infected cells were also quickly frozen in liquid helium and observed under the electron microscope as described previously (Heuser, 1989).

Virus attachment to infected and uninfected chlorella cells

Fifteen milliliters of chlorella cells (1.5×10^7 cells/ml) were infected with PBCV-1 at a m.o.i. of 5, incubated for 15 min at 25°C, and divided into 1.5-ml samples. At various times after the initial PBCV-1 infection, virus P31 [an antigenic variant of PBCV-1 (Wang *et al.*, 1993)] was added at a m.o.i. of 5 and incubated for 15 min. Samples were treated with PBCV-1 antiserum for 15 min, followed by low-speed centrifugation to remove algae, attached virus, and unattached PBCV-1 virus complexed to antibody. The supernatant was titered for unattached P31.

Northern and Southern analyses

Chlorella cells (1×10^9) were collected at various times after PBCV-1 infection, frozen in liquid nitrogen, and stored at −80°C. RNA was extracted using the Trizol reagent (GIBCO BRL, Gaithersburg, MD), electrophoresed under denaturing conditions on 1.5% agarose/formaldehyde gels, stained with ethidium bromide, and transferred to nylon membranes. Membranes were subsequently photographed under UV illumination to visualize transferred RNA. The RNA was hybridized with either *has*, *gfat*, or *udp-glcdh* specific probes labeled with ^{32}P using a random primed DNA labeling kit (GIBCO BRL) at 65°C in 50 mM $NaPO_4$, 1% BSA, and 2% SDS. After hybridization, radioactivity bound to the membranes was detected and quantified using a Storm 840 PhosphorImager and ImageQuant software (Molecular Dynamics, Inc., Sunnyvale, CA). To account for loading differences between samples, the relative amount of the 3.6-kb rRNA in each lane was determined by converting the photographs of the stained membranes to digital images using a Hewlett Packard ScanJet 4C scanner and analyzing the images using the ImageQuant software.

Chlorella virus DNAs for dot blots were denatured and applied to nylon membranes (Micron Separation Inc., Westborough, MA), fixed by UV cross-linking, and hybridized with the same *has* gene probe used for the Northern analyses. Radioactivity bound to the filters was detected as described above.

Other procedures

DNA fragments were sequenced from both strands at the University of Nebraska-Lincoln Center for Biotechnology DNA sequencing core facility. DNA and protein sequences were analyzed with the University of Wisconsin Genetics Computer Group package of programs (Genetics Computer Group, 1997). The GenBank Accession numbers for the *has* genes from viruses AL-2C, CA-4A, MA-1E, NC-1C, and XZ-5C are AF113753, AF113754, AF113755, AF113756, and AF113757, respectively.

ACKNOWLEDGMENTS

We thank Mike Nelson and Les Lane for helpful discussions. We also thank Martin Dickman and Jeff Rollins for the use of and assistance with the Zeiss microscope. This investigation was supported in part by U.S. Public Health Service Grants GM-32441 (J.V.E.) and GM-56497 (P.L.D.), NSF-EPSCoR Cooperative Agreement EPS-9255225 (J.V.E.), and an AOC grant from the University of Nebraska Biotechnology Center (J.V.E.).

REFERENCES

Aruffo, A., Stamenkovic, I., Melnick, M., Underhill, C. B., and Seed, B. (1990). CD44 is the principal cell surface receptor for hyaluronate. *Cell* 61,1303–1313.

Carter, G. R., and Annau, E. (1953). Isolation of capsular polysaccharides from colonial variants of *Pasteurella multocida*. *Am. J. Vet. Res.* 14, 475–478.

Chase, T. E., Nelson, J. A., Burbank, D. E., and Van Etten, J.L. (1989). Mutual exclusion occurs in a chlorella-like green alga inoculated with two viruses. *J. Gen. Virol.* 70, 1829–1836.

Culty, M., Miyake, K., Kincade, P. W., Silorsky, E., Butcher, E. C., and Underhill, C. (1990). The hyaluronate receptor is a member of the CD44 (H-CAM) family of cell surface glycoproteins. *J. Cell Biol.* 111, 2765–2774.

DeAngelis, P. L., Jing, W., Graves, M. V., Burbank, D. E., and Van Etten, J. L. (1997). Hyaluronan synthase of chlorella virus PBCV-1. *Science* 278, 1800–1803.

Genetics Computer Group. (1997). Wisconsin Package Version 9.1, Madison, WI.

Hall, C. L., Yang, B., Yang, X., Zhang, S., Turley, M., Samuel, S., Lange, L. A., Wang, C., Curpen, G. D., Savani, R. C., Greenberg, A. H., and Turley, E. A. (1995). Overexpression of the hyaluronan receptor RHAMM is transforming and is also required for H-*ras* transformation. *Cell* 82, 19–28.

Hardwick, C., Hoare, K., Owens, R., Hohn, H. P., Hook, M., Moore, D., Cripps, V., Austen, L., Nance, D. M., and Turley, E. A. (1992). Molecular cloning of a novel hyaluronan receptor that mediates tumor cell motility. *J. Cell Biol.* 117, 1343–1350.

Hascall, V. C., Calabro, A., Midura, R. J., and Yanagishita, M. (1994). Isolation and characterization of proteoglycans. *Methods Enzymol.* 230, 390–417.

Heuser, J. E. (1989). A detailed protocol for deep-etch electron microscopy of molecules. *J. Electron. Microsc. Methods* 13, 244–263.

Ho, C. K., Van Etten, J. L., and Shuman, S. (1996). Expression and characterization of an RNA capping enzyme encoded by chlorella virus PBCV-1. *J. Virol.* 70, 6658–6664.

Husmann, L. K., Yung, D. L., Hollingshead, S. K., and Scott, J. R. (1997). Role of putative virulence factors of *Streptococcus pyogenes* in mouse models of long-term throat colonization and pneumonia. *Infect. Immunol.* 65, 1422–1430.

Jaworski, D. M., Kelly, G. M., and Hockfield, S. (1994). BEHAB, a new member of the proteoglycan tandem repeat family of hyaluronan-binding proteins that is restricted to the brain. *J. Cell Biol.* 125, 495–509.

Kass, E. H., and Seastone, C. V. (1944). The role of the mucoid polysaccharide (hyaluronic acid) in the virulence of group A hemolytic streptococci. *J. Exp. Med.* 79, 319–330.

Kutish, G. F., Li, Y., Lu, Z., Furuta, M., Rock, D. L., and Van Etten, J.L. (1996). Analysis of 76 kb of the chlorella virus PBCV-1 330-kb genome: Map positions 182 to 258. *Virology* 223, 303–317.

Landstein, D., Graves, M. V., Burbank, D. E., DeAngelis, P. L., and Van Etten, J. L. (1998). Chlorella virus PBCV-1 encodes functional glutamine:fructose-6-phosphate amidotransferase and UDP-glucose dehydrogenase enzymes. *Virology* 250, 388–396.

Laurent, T. C., and Fraser, J. R.E. (1992). Hyaluronan. *FASEB J.* 6, 2397–2404.

Li, Y., Lu, Z., Burbank, D. E., Kutish, G. F., Rock, D. L., and Van Etten, J. L. (1995). Analysis of 43 kb of the chlorella virus PBCV-1 330-kb genome: Map position 45 to 88. *Virology* 212, 134–150.

Li, Y., Lu, Z., Sun, L., Ropp, S., Kutish, G. F., Rock, D. L., and Van Etten, J.L. (1997). Analysis of 74 kb of DNA located at the right end of the 330-kb chlorella virus PBCV-1 genome. *Virology* 237, 360–377.

Lu, Z., Li, Y., Que, Q., Kutish, G. F., Rock, D. L., and Van Etten, J. L. (1998). Analysis of 94 kb of the chlorella virus PBCV-1 330-kb genome: Map positions 88 to 182. *Virology* 216, 102–123.

Lu, Z., Li, Y., Zhang, Y., Kutish, G. F., Rock, D. L., and Van Etten, J. L. (1995). Analysis of 45 kb of DNA located at the left end of the chlorella virus PBCV-1 genome. *Virology* 206, 339–352.

Meints, R. H., and Pardy, R. L. (1980). Quantitative demonstration of cell surface involvement in a plant-animal symbiosis: Lectin inhibition of reassociation. *J. Cell Sci.* 43, 239–251.

Miyake, K., Underhill, C. B., Lesley, J., and Kincade, P. W. (1990). Hyaluronate can function as a cell adhesion molecule and CD44 participates in hyaluronate recognition. *J. Exp. Med.* 172, 69–75.

O'Reilly, D. R. (1995). Baculovirus-encoded ecdysteroid UDP-glucosyltransferases. *Insect Biochem. Mol. Biol.* 25, 541–550.

Philipson, L. H., and Schwartz, N. B. (1984). Subcellular localization of hyaluronate synthetase in oligodendroglioma cells. *J. Biol. Chem.* 259, 5017–5023.

Pool, R. R. (1979). The role of algal antigenic determinants in the recognition of potential algal symbionts by cells of chlorohydra. *J. Cell Sci.* 35, 367–379.

Que, Q., Li, Y., Wang, I. N., Lane, L. C., Chaney, W. G., and Van Etten, J. L. (1994). Protein glycosylation and myristylation in chlorella virus PBCV-1 and its antigenic variants. *Virology* 203, 320–327.

Reisser, W. (1992). Basic mechanisms of signal exchange, recognition, specificity, and regulation in endosymbiotic systems. In "Algae and Symbioses" (W. Reisser, Ed.). Biopress Limited, Bristol, UK.

Reisser, W., Burbank, D. E., Meints, R. H., Becker, B., and Van Etten, J.L. (1991). Viruses distinguish symbiotic *Chlorella* sp. of *Paramecium bursaria*. *Endocytobiosis Cell Res.* 7, 245–251.

Reisser, W., Burbank, D. E., Meints, S. M., Meints, R. H., Becker, B., and Van Etten, J.L. (1988). A comparison of viruses infecting two different chlorella-like green algae. *Virology* 167, 143–149.

Renesto, P., Sansonetti, P. J., and Guillen, N. (1997). Interaction between *Entamoeba histolytica* and intestinal epithelial cells involves a CD44 cross-reactive protein expressed on the parasite surface. *Infect. Immunol.* 65, 4330–4333.

Salustri, A., Ulisse, S., Yanagishita, M., and Hascall, V. C. (1990). Hyaluronic acid synthesis by mural granulosa cells and cumulus cells in vitro is selectively stimulated by a factor produced by oocytes and by transforming growth factor-β. *J. Biol. Chem.* 265, 19517–19523.

Schmidt, K. H., Gunther, E., and Courtney, H. S. (1996). Expression of both M protein and hyaluronic acid capsule by group A streptococcal strains results in a high virulence for chicken embryos. *Med. Microbiol. Immunol.* 184, 169–173.

Sherman, L., Sleeman, J., Herrlich, P., and Ponta, H. (1994). Hyaluronate receptors: Key players in growth, differentiation, migration and tumor progression. *Curr. Opin. Cell Biol.* 6, 726–733.

Tengblad, A. (1980). Quantitative analysis of hyaluronate in nanogram amounts. *Biochem. J.* 185, 101–105.

Van Etten, J. L., Burbank, D. E., Joshi, J., and Meints, R. H. (1984). DNA synthesis in a chlorella-like alga following infection with the virus PBCV-1. *Virology* 134, 443–449.

Van Etten, J. L., Burbank, D. E., Kuczmarski, D., and Meints, R. H. (1983a). Virus infection of culturable chlorella-like algae and development of a plaque assay. *Science* 219, 994–996.

Van Etten, J. L., Burbank, D. E., Schuster, A. M., and Meints, R. H. (1985). Lytic viruses infecting a chlorella-like alga. *Virology* 140, 135–143.

Van Etten, J. L., Burbank, D. E., Xia, Y., and Meints, R.H. (1983). Growth cycle of a virus, PBCV-1, that infects chlorella-like algae. *Virology* 126, 117–125.

Van Etten, J. L., Lane, L. C., and Meints, R.H. (1991). Viruses and virus-like particles of eukaryotic algae. *Microbiol. Rev.* 55, 586–620.

Van Etten, J. L., Meints, R. H., Burbank, D. E., Kuczmarski, D., Cuppels, D. A., and Lane, L.C. (1981). Isolation and characterization of a virus from the intracellular green alga symbiotic with *Hydra viridis*. *Virology* 113, 704–711.

Wang, I. N., Li, Y., Que, Q., Bhattacharya, M., Lane, L. C., Chaney, W. G., and Van Etten, J. L. (1993). Evidence for virus-encoded glycosylation specificity. *Proc. Natl. Acad. Sci. USA* 90, 3840–3844.

West, D. C., Hampson, I. N., Arnold, F., and Kumar, S. (1985). Angiogenesis induced by degradation products of hyaluronic acid. *Science* 228, 1324–1326.

Yamada, T., Higashiyama, T., and Fukuda, T. (1991). Screening of natural waters for viruses which infect chlorella cells. *Appl. Environ. Microbiol.* 57, 3433–3437.

supposition in a native ecosystem. The data presented here demonstrating that herbivory by Serengeti grazers tangibly accelerates the mineralization of two minerals of considerable importance in animal nutrition are consistent with simulation results from grassland ecosystem models (*3*). In addition, they indicate that the accelerated recycling of plant-available Na is probably the mechanism leading to levels of that animal nutrient in grazer-exploited Serengeti grasslands that are sufficient to alleviate nutritional shortage in the grazers, particularly reproductive females and growing young (*5*).

Mammalian herbivores have been pervasive in grasslands through evolutionary time (*12*), their levels of forage consumption are considerable (*6, 13*), and the animals accelerate rather than retard nutrient cycling. The intensity of the plant-herbivore interaction in grasslands, and its evolutionary antiquity, may have attenuated detrimental interaction effects through coevolution (*1*). Overgrazing of grasslands, on the other hand, which is commonly associated with the replacement of free-ranging wild herbivores with livestock and the resulting higher animal densities (*14*), often causes the replacement of highly palatable forages (*15*) that produce easily decomposable litter (*10*) with other plant species of lower nutritional quality and decomposability.

These data provide evidence that a terrestrial grazer can modify ecosystem processes in such a way as to alleviate nutritional deficiencies and, therefore, plausibly to elevate the carrying capacity of the ecosystem. The data also identify accelerated nutrient cycling as an important property of habitats that are critical to large mammal conservation (*16*). The coupling of animal site preference with nutritional effects could provide a guide for identifying sites essential for planning large mammal conservation in natural ecosystems. In addition, the presence of such sites, and the role of mammals in maintaining them, provide clear evidence that habitat deterioration is not an inescapable consequence of increased density of organisms (*1*).

REFERENCES AND NOTES

1. A. F. Hunter and L. W. Aarssen, *BioScience* 38, 34 (1988); T. S. Bianchi, C. G. Jones, M. Shachak, *Trends Ecol. Evol.* 4, 234 (1989); M. D. Bertness and R. Callaway, *ibid.* 9, 191 (1994); C. G. Jones, J. H. Lawton, M. Shachak, *Oikos* 69, 373 (1994).
2. R. G. Woodmansee, *BioScience* 28, 448 (1978); S. J. McNaughton, *Oikos* 40, 329 (1983); J. K. Detling, in *Concepts of Ecosystem Ecology: A Comparative View*, L. R. Pomeroy and J. J. Alberts, Eds. (Springer-Verlag, New York, 1988), pp. 131–148; D. A. Frank *et al.*, *Biogeochemistry* 26, 163 (1994).
3. E. A. Holland and J. K. Detling, *Ecology* 71, 1040 (1990); E. A. Holland, W. J. Parton, J. K. Detling, D. L. Coppock, *Am. Nat.* 140, 685 (1992); S. W. Seagle, S. J. McNaughton, R. W. Ruess, *Ecology* 73, 1105 (1992); D. A. Frank and R. D. Evans, *ibid.* 78, 2238 (1997).
4. J. Pastor, R. J. Naiman, B. Dewey, P. F. McInes, *BioScience* 38, 770 (1988); J. Pastor, B. Dewey, R. J. Naiman, P. F. McInes, Y. Cohen, *Ecology* 74, 467 (1993).
5. S. J. McNaughton, *Nature* 334, 343 (1988); *ibid.* 345, 613 (1990).
6. ———, *Ecol. Monogr.* 53, 291 (1983); *ibid.* 55, 259 (1985).
7. Procedures followed R. S. Siegel, *Soil Sci. Soc. Am. J.* 44, 943 (1980); R. J. Raison, M. J. Connell, P. K. Khanna, *Soil Biol. Biochem.* 19, 521 (1987); A. Mehlich, *Comm. Soil Sci. Plant Anal.* 15, 1409 (1984); F. J. Adamsen, D. S. Bigelow, G. R. Scott, *ibid.* 16, 833 (1985); and (*5*). Soils for N analyses were extracted with 1 M KCl, soils for Na analyses were extracted with Mehlich solution No. 3, and both extracts were preserved with phenylmercuric acetate. Mineral N as NH_4 and NO_3 was determined by continous flow analysis (Lachat QuikChem AE, Milwaukee, WI), and Na was determined by inductively coupled argon plasma spectrometry (Leman Labs PS3000 (Hudson, MA) in simultaneous mode). Data were analyzed by nested (plot within site) repeated measures of analysis of variance (C. N. von Ende, in *Design and Analysis of Ecological Experiments*, S. M. Scheiner and J. Gurevitch, Eds. (Chapman and Hall, New York, 1993), pp. 113–137).
8. J. K. Detling and E. L. Painter, *Oecologia* 57, 65 (1983); W. G. Gold and M. M. Caldwell, *ibid.* 81, 437 (1989).
9. K. Mengel and E. A. Kirkby, *Principles of Plant Nutrition* (International Potash Institute, Bern, Switzerland, 1982), pp. 61–62.
10. D. A. Wedin and D. Tilman, *Oecologia* 84, 433 (1990); S. Enriquez, C. M. Duarte, K. Sand-Jensen, *ibid.* 94, 457 (1993); D. A. Wedin and D. Tilman, *Science* 274, 1720 (1996); J. P. Grime, J. H. C. Cornelissen, K. Thompson, J. G. Hodgson, *Oikos* 77, 489 (1996).
11. S. J. McNaughton, F. F. Banyikwa, M. M. McNaughton, *Ecosystems*, in press.
12. S. J. McNaughton, in *Plant-Animal Interactions: Evolutionary Ecology in Tropical and Temperate Regions*, P. W. Price, T. M. Lewinsohn, G. W. Fernandes, W. W. Benson, Eds. (Wiley-Interscience, New York, 1991), pp. 509–522.
13. S. J. McNaughton, M. Oesterheld, D. A. Frank, K. J. Williams, *Nature* 341, 142 (1989); S. J. McNaughton, in *Terrestrial Global Productivity: Past, Present, Future*, J. Roy, B. Schmid, H. A. Mooney, Eds. (Academic Press, San Diego, CA, in press).
14. M. Oesterheld, O. Sala, S. J. McNaughton, *Nature* 356, 234 (1992).
15. R. F. Noss and A. Y. Cooperider, *Saving Nature's Legacy* (Island Press, Washington, DC, 1994).
16. S. J. McNaughton and F. F. Banyikwa, in *Serengeti II: Dynamics, Management, and Conservation of an Ecosystem*, A. R. E. Sinclair and P. Arcese, Eds. (Univ. of Chicago Press, Chicago, IL, 1995), pp. 49–70.

17. Research was supported by NSF grant DEB-9312435. Permission to do research and live in Serengeti National Park was granted by the Scientific Research Council of Tanzania, Tanzania National Parks, Serengeti National Park, and the Serengeti Wildlife Research Institute. Housing and laboratory space were provided by the Serengeti Wildlife Research Centre. We are grateful to E. M. Peter for tireless field and lab assistance.

8 August 1997; accepted 28 October 1997

Hyaluronan Synthase of Chlorella Virus PBCV-1

Paul L. DeAngelis,* Wei Jing, Michael V. Graves, Dwight E. Burbank, James L. Van Etten

Sequence analysis of the 330-kilobase genome of the virus PBCV-1 that infects a chlorella-like green algae revealed an open reading frame, A98R, with similarity to several hyaluronan synthases. Hyaluronan is an essential polysaccharide found in higher animals as well as in a few pathogenic bacteria. Expression of the *A98R* gene product in *Escherichia coli* indicated that the recombinant protein is an authentic hyaluronan synthase. *A98R* is expressed early in PBCV-1 infection and hyaluronan is produced in infected algae. These results demonstrate that a virus can encode an enzyme capable of synthesizing a carbohydrate polymer and that hyaluronan exists outside of animals and their pathogens.

Hyaluronan or hyaluronic acid (HA), a member of the glycosaminoglycan family that also includes heparin and chondroitan, is a linear polysaccharide composed of alternating β1,4-glucuronic acid (β1,4-GlcA) and β1,3-*N*-acetylglucosamine (β1,3-GlcNAc) groups. Typically the full-length polymer chains are composed of 10^3 to 10^4 monosaccharides (10^6 to 10^7 daltons). HA is an important structural element in the vitreous humor of eye, synovial fluid, and skin of vertebrates (*1*). Furthermore, HA interacts with proteins such as CD44, RHAMM, and fibrinogen, thereby influencing many natural processes such as angiogenesis, cancer, cell motility, wound healing, and cell adhesion (*2*). HA also constitutes the extracellular capsules of certain bacterial pathogens such as group A and C *Streptococcus* and *Pasteurella multocida* type A (*3, 4*). These capsules act as virulence factors that protect the microbes from phagocytosis and complement during infection (*5, 6*). Because HA, a component of the host tissues, is not normally immunogenic, the capsule serves as molecular camouflage (*7*).

HA synthases (HASs) are integral mem-

P. L. DeAngelis and W. Jing, Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center, 940 Stanton L. Young Boulevard, Oklahoma City, OK 73104, USA.
M. V. Graves, D. E. Burbank, J. L. Van Etten, Department of Plant Pathology, University of Nebraska, Lincoln, NE 68583-0722, USA.

*To whom correspondence should be addressed. E-mail: paul-deangelis@OUHSC.edu

brane proteins that polymerize the HA molecule using activated uridine diphosphate (UDP)–sugar nucleotides as substrates. Amino acid sequences for some HASs have been deduced from gene sequencing (*8*); their sizes range from 419 to 588 residues. The vertebrate enzymes (DG42, HAS1, HAS2, and HAS3) and streptococcal HasA have several regions of sequence similarity. Recently, while sequencing the double-stranded DNA genome of virus PBCV-1 (*Paramecium bursaria* chlorella virus), we unexpectedly discovered an open reading frame (ORF), A98R (GenBank accession number U42580), encoding a 568-residue protein with similarity to the known HASs (28 to 33% amino acid identity in pairwise comparisons by FASTA) (Fig. 1).

PBCV-1 is the prototype of a family (Phycodnaviridae) of large (175 to 190 nm in diameter) polyhedral, plaque-forming viruses that replicate in certain unicellular, eukaryotic chlorella-like green algae (*9*). PBCV-1 virions contain at least 50 different proteins and a lipid component located inside the outer glycoprotein capsid (*10*). The PBCV-1 genome is a linear, nonpermuted 330-kb double-stranded DNA molecule with covalently closed hairpin ends (*11*).

On the basis of its deduced amino acid sequence, the A98R gene product should be an integral membrane protein. To test this hypothesis, we produced recombinant A98R protein in *Escherichia coli* and assayed the membrane fraction for HAS activity (*12, 13*). UDP-GlcA and UDP-GlcNAc were incorporated into polysaccharide by the membrane fraction derived from cells containing the A98R ORF on a plasmid, pCVHAS, (average specific activity of 2.5 pmol of GlcA transferred per minute per microgram of protein), but not by samples from control cells (<0.001 pmol of GlcA transferred per minute per microgram of protein). No activity was detected in the soluble fraction of cells transformed with pCVHAS. UDP-GlcA and UDP-GlcNAc were simultaneously required for polymerization. The activity was optimal in Hepes buffer at pH 7.2 in the presence of 15 mM $MnCl_2$, whereas no activity was detected if the metal ion was omitted. The ions Mg^{2+} and Co^{2+} were ~20% as effective as Mn^{2+} at similar concentrations. The *P. multocida* HAS (*14*) has a similar metal requirement, but other HASs prefer Mg^{2+}.

We also tested the specificity of recombinant A98R for UDP-sugars (*15*). Only the two authentic HA precursors were incorporated into polysaccharide; neither UDP-galacturonic acid (UDP-GalA) nor UDP-*N*-acetylgalactosamine (UDP-GalNAc), the C4 epimers of UDP-GlcA or UDP-GlcNAc, respectively, were incorporated. Likewise, UDP-glucose (UDP-Glc) was not polymerized in place of either HA precursor. This strong substrate specificity for UDP-GlcA and UDP-GlcNAc is a general feature of the HASs HasA (*13*) and DG42 (*16*).

The recombinant A98R enzyme synthesized a polysaccharide with an average molecular size of 3×10^6 to 6×10^6 daltons (Fig. 2), which is smaller than that of the HA synthesized by recombinant HasA or DG42 in vitro (~10^7 daltons and ~5×10^6 to 8×10^6 daltons, respectively) (*13, 16*).

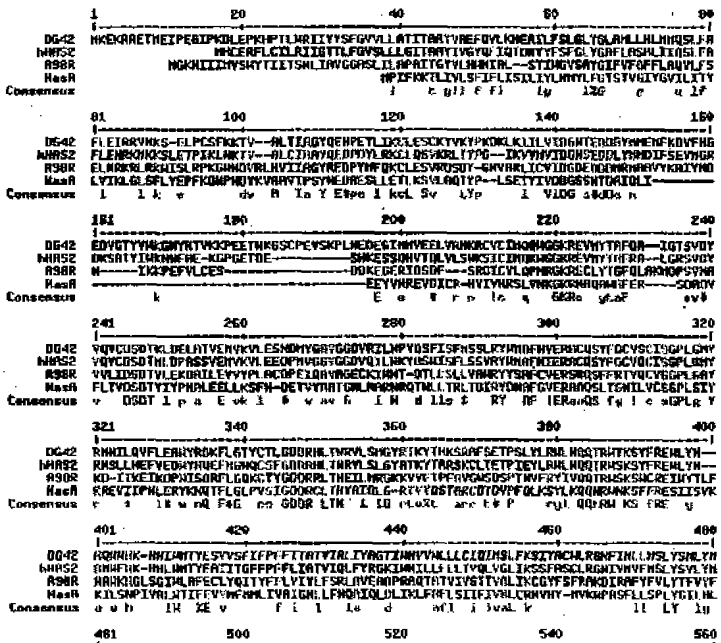


Fig. 1. Sequence similarity of HASs. The Multalin program (*26*) was used to align the amino acid sequences of HASs *Xenopus laevis* DG42, human HAS2, PBCV-1 A98R, and *Streptococcus pyogenes* HasA (red; 90% consensus; green, 50% consensus, as calculated by Multalin) (*8*). In the consensus sequence, the symbols are: !, any one of I or V; $, any one of L or M; %, any one of F or Y; #, any one of N,D,E, or Q. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.

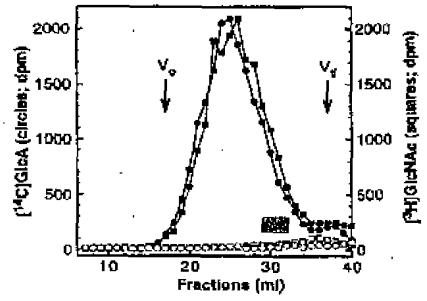


Fig. 2. Size exclusion chromatography of polymer product of recombinant A98R HAS. Membranes derived from *E. coli* cells transformed with pCVHAS were incubated with both radiolabeled HA precursors diluted to the same specific activity (*27*). After deproteinization and removal of unincorporated precursors, samples were injected onto a Sephacryl S-500HR size exclusion column, and the radioactivity in the fractions was measured (^{3}H, solid squares; ^{14}C, solid circles). A duplicate sample was treated with HA lyase before deproteinization and chromatography (^{3}H, open squares; ^{14}C, open circles); no polymer remains after digestion. Size standards: V_o arrow, void volume, HA derived from recombinant streptococcal HasA (17 ml; ≥2 × 10^7 daltons) (*13*); crosshatched box, blue dextran (29 to 32 ml; average molecular size 2 × 10^6 daltons; Pharmacia); V_t arrow, totally included volume, UDP-sugars (37 ml).

The polysaccharide was completely degraded by *Streptomyces hyalurolyticus* HA lyase, an enzyme that depolymerizes HA but not structurally related glycosaminoglycans such as heparin and chondroitan (*17*).

We examined PBCV-1–infected chlorella cells for *A98R* gene expression. A ~1700-nucleotide A98R transcript appeared about 15 min after infection and disappeared by 60 min after infection (*18*), indicating that *A98R* is an early gene. Consequently, we assayed membrane fractions from uninfected and PBCV-1–infected chlorella cells at 50 and 90 min after infection for HAS activity. Infected cells, but not uninfected cells, had activity (Table 1). Like the bacterially derived recombinant A98R enzyme, radioactive label incorporation from UDP-[^{14}C]GlcA into polysaccharide depended on both Mn^{2+} and UDP-GlcNAc. This labeled product was also degraded by HA lyase. Disrupted PBCV-1 virions had no HAS activity.

PBCV-1–infected chlorella cells were analyzed for HA polysaccharide by means of a highly specific ^{125}I-labeled HA-binding protein (*19*, *20*). Extracts from cells at 50 and 90 min after infection contained substantial amounts of HA (0.7 and 1400 ng per microgram of protein, respectively), but not extracts from uninfected algae (<0.04 ng per microgram of protein) or disrupted PBCV-1 virions (<0.04 ng per microgram of dry weight). The labeled HA-binding protein also interacted with intact infected cells at 50 and 90 min after infection, but not with healthy cells (*21*). Therefore, a considerable portion of the newly synthesized HA polysaccharide was immobilized at the outer cell surface of the infected algae.

Table 1. HAS activity of membranes derived from *Chlorella* cells infected with PBCV-1. The membrane fractions (370 μg of protein) from uninfected cells or cells at 50 and 90 min after infection (a.i.) were assayed with UDP-[^{14}C]GlcA (60 μM, 0.02 μCi) in parallel reactions containing the following components as indicated (300 μM UDP-GlcNAc or 15 mM $MnCl_2$ or both) for 1 hour at 30°C (*28*). HAS specific activity (presented as picomoles of [^{14}C]GlcA transferred per hour per milligram of protein) was detected in the algal membranes after infection with PBCV-1, but not in uninfected cells.

Sample	UDP-GlcNAc	Mn^{2+}	HAS specific activity
Uninfected	+	+	≤6
	+	−	≤6
	−	+	≤6
50 min a.i.	+	+	42
	+	−	≤6
	−	+	≤6
90 min a.i.	+	+	170
	+	−	≤6
	−	+	≤6

The extracellular HA does not play any obvious role in the interaction between the virus and its algal host because neither plaque size nor plaque number was altered by including either testicular hyaluronidase (465 units/ml) or free HA polysaccharide (100 μg/ml) in the top agar of the PBCV-1 plaque assay (*9*).

Among chlorella viruses, HA biosynthesis during infection is not limited to the PBCV-1 prototype strain. Thirty-three independently isolated and plaque-purified viruses from the United States, South America, Asia, and Australia were tested for the presence of an A98R-like gene and for the ability to direct production of HA polysaccharide in *Chlorella* NC64A. Dot-blot hybridization analyses of the individual viral genomes with the PBCV-1 A98R probe indicated that 19 isolates (58%) had a similar gene; the algal host DNA did not cross-react with the probe (*21*). Chlorella cells infected with each of these 19 viruses produced cell surface HA as measured by interaction with the ^{125}I-HA–binding protein (*21*).

Surprisingly, the PBCV-1 genome also has additional genes, named *A609L* and *A100R*, that encode for a UDP-Glc dehydrogenase (UDP-Glc DH) and a glutamine: fructose-6-phosphate amidotransferase (GFAT), respectively. UDP-Glc DH converts UDP-Glc into UDP-GlcA, a required precursor for HA biosynthesis. GFAT converts fructose-6-phosphate into glucosamine-6-phosphate, an intermediate in the UDP-GlcNAc metabolic pathway. Both of these PBCV-1 genes, like the *A98R* HAS, are expressed early in infection and encode enzymatically active proteins (*22*); however, these three genes do not function as an operon. Although two of these genes, *A98R* and *A100R*, are near one another in the viral genome (bases 50,901 to 52,607 and 52,706 to 54,493, respectively), *A609L* is located ~240 kb away and is transcribed in the opposite orientation (bases 292,916 to 291,747). The presence of multiple enzymes in the HA biosynthesis pathway indicates that HA production must serve an important function in the life cycle of these chlorella viruses.

The details of the natural history of the phycodnaviruses are unknown. These viruses are ubiquitous in freshwater collected worldwide, and titers as high as 4×10^4 infectious viruses per milliliter of native water have been reported (*23*). The only known hosts for these viruses are chlorella-like green algae, which normally live as hereditary endosymbionts in some isolates of the ciliate, *P. bursaria*. In the symbiotic unit, algae are enclosed individually in perialgal vacuoles and are surrounded by a host-derived membrane (*24*). The endosymbiotic chlorella are resistant to virus infection and are only infected when they are outside the paramecium (*9*). We hypothesize that HA synthesis and its accumulation on the algal surface may block the uptake of virus-infected algae by the paramecium. Alternatively, the chlorella viruses might have another host in nature (such as an aquatic animal); perhaps the virus is transmitted because this other host is attracted to or binds to the HA polysaccharide on virus-infected algae.

As depicted in Fig. 1, HASs of *Streptococcus*, vertebrates, and PBCV-1 have many motifs of two to four residues that occur in the same relative order. These conserved motifs probably reflect domains crucial for HA biosynthesis. Regions of similarity between HASs and other enzymes that synthesize β-linked polysaccharides from UDP-sugar precursors are also being discovered as more glycosyltransferases are sequenced (*25*). The significance of these similar structural motifs will become more apparent as the three-dimensional structures of glycosyltransferases are determined.

The fact that *Chlorella* virus PBCV-1 encodes a functional glycosyltransferase that can synthesize HA is contrary to the general observation that viruses either (i) use host cell glycosyltransferases to create new carbohydrate structures, or (ii) accumulate host cell glycoconjugates during virion maturation. Furthermore, HA has been generally regarded as restricted to animals and a few of their virulent bacterial pathogens. Though many plant carbohydrates have been characterized, to our knowledge, neither HA nor a related analog has previously been detected in cells of plants or protists.

REFERENCES AND NOTES

1. T. C. Laurent and J. R. E. Fraser, *FASEB J.* 6, 2397 (1992).
2. C. B. Knudson and W. Knudson, *ibid.* 7, 1234 (1993).
3. E. H. Kass and C. V. Seastone, *J. Exp. Med.* 79, 319 (1944).
4. G. R. Carter and E. Anneu, *Am. J. Vet. Res.* 14, 475 (1953).
5. B. G. Harmon, J. R. Glisson, K. S. Latimer, W. L. Steffens, J. C. Nunnally, *ibid.* 52, 1507 (1991).
6. M. R. Wessels, A. E. Moses, J. B. Goldberg, T. J. DiCesare, *Proc. Natl. Acad. Sci. U.S.A.* 88, 8317 (1991).
7. A. W. Quinn and K. P. Sing, *Proc. Soc. Exp. Biol. Med.* 95, 290 (1957).
8. P. L. DeAngelis, J. Papaconstantinou, P. H. Weigel, *J. Biol. Chem.* 268, 19181 (1993); P. H. Weigel, V. C. Hascall, M. Tammi, *ibid.* 272, 13997 (1997).
9. J. L. Van Etten, L. C. Lane, R. H. Meints, *Microbiol. Rev.* 55, 586 (1991); J. L. Van Etten, *Mol. Cells* 5, 99 (1995); ——— D. E. Burbank, D. Kuczmarski, R. H. Meints, *Science* 219, 994 (1983).
10. M. P. Skadla, D. E. Burbank, Y. Xia, R. H. Meints, J. L. Van Etten, *Virology* 135, 308 (1984); I.-N. Wang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 90, 3640 (1993).
11. J. Rohozinski, L. E. Girton, J. L. Van Etten, *Virology*

168, 363 (1989); Y. Zhang, P. Strasser, R. Grabherr, J. L. Van Etten, *ibid.* 202, 1079 (1994).
12. The A98R ORF was cloned after amplification of genomic viral DNA with 22 cycles of polymerase chain reaction (PCR) with Taq polymerase [F. M. Ausübel *et al.*, in *Short Protocols in Molecular Biology* (Wiley, New York, ed. 3, 1995)]. The oligonucleotide primers [sense 5′-gagagctccATGGGTAAAAA-CATTATCATAATGG-3′; antisense 5′-gcatgtcgacT-CACACAGACTGAGCATTGGTAG-3′; Great American Gene Company] contained Nco I or Sal I restriction sites (underlined, respectively) flanking the ORF encoding 568 amino acids (uppercase letters). Codons 4 and 5 were altered to optimize bacterial expression. The PCR product was purified, digested with Nco I and Sal I, and ligated into a modified version of the plasmid pET-8C [B. A. Moffatt and F. W. Studier, *J. Mol. Biol.* 189, 113 (1986)] (it has an extra Sal I site in the polylinker) cleaved with Nco I and partially digested with Sal I. This construct placed the A98R ORF under the control of a T7 phage promoter. The resulting plasmid, pCVHAS, was transformed into the expression host, *E. coli* BL21(DE3). The A98R protein was expressed by induction with 1 mM isopropylthiogalactoside. After 3 to 5 hours of further growth, the membrane fraction was isolated (*13*). Control membrane preparations were made from cultures with the same vector containing an irrelevant gene (a protein kinase). Total protein was measured according to M. M. Bradford [*Anal. Biochem.* 72, 248 (1976)]. The paper chromatography method was used to assay for HAS activity (*13*).
13. P. L. DeAngelis and P. H. Weigel, *Biochemistry* 33, 9033 (1994).
14. P. L. DeAngelis, *ibid.* 35, 9768 (1996).
15. Assays with *E. coli*–derived A98R were incubated either with 120 μM UDP-[^{14}C]GlcA or with 150 μM UDP-[^{3}H]GlcNAc. The reactions also contained one unlabeled sugar nucleotide (an authentic precursor or UDP-Glc, UDP-GalA, or UDP-GalNAc) at 300 μM. Less than 5% of the maximal incorporation (assay with UDP-GlcA and UDP-GlcNAc present) was detected if an unnatural UDP-sugar was substituted for UDP-GlcA or UDP-GlcNAc, or if only a single precursor was present.
16. P. E. Pummill, A. M. Achyuthan, P. L. DeAngelis, in preparation.
17. T. Ohya and Y. Kaneko, *Biochim. Biophys. Acta* 198, 607 (1970).
18. M. V. Graves, D. Landstein, J. L. Van Etten, unpublished results.
19. A. Tengblad, *Biochem. J.* 185, 101 (1980).
20. A competitive radiometric assay based on a ^{125}I-labeled HA-binding protein (*19*) (Pharmacia HA Test) was used to measure the amount of HA in disrupted (freeze-thawed) virus particles or the cultures of NC64A cells. The cells were disrupted by vigorous agitation with glass beads (1-mm beads, agitated for 3 min four times; Biospec Mini-Beadbeater-8). The cell lysate was clarified by centrifugation (15,000*g*, 5 min) before assaying (average of two determinations).
21. P. L. DeAngelis, M. V. Graves, J. L. Van Etten, unpublished results.
22. D. Landstein, P. L. DeAngelis, J. L. Van Etten, in preparation.
23. J. L. Van Etten, D. E. Burbank, A. M. Schuster, R. H. Meints, *Virology* 140, 135 (1985); T. Yamada, T. Higashiyama, T. Fukuda, *Appl. Environ. Microbiol.* 57, 3433 (1991).
24. W. Reisser, Ed., *Algae and Symbioses* (Biopress, Bristol, UK, 1992).
25. P. L. DeAngelis, N. Yang, P. H. Weigel, *Biochem. Biophys. Res. Commun.* 199, 1 (1994); I. M. Saxena *et al.*, *J. Bacteriology* 177, 1419 (1995).
26. F. Corpet, *Nucleic Acids Res.* 16, 10881 (1988).
27. Membranes (660 μg of protein) were incubated with 120 μM UDP-[^{14}C]GlcA (0.36 μCi) and 840 μM UDP-[^{3}H]GlcNAc (2.6 μCi) in 300 μl of 50 mM Hepes, pH 7.2, with 15 mM $MnCl_2$ for 3 hours at 30°C. EDTA (18 mM final concentration) was then added to stop the HAS activity. Half of the reaction was deproteinized by treatment with 0.5% SDS (w/v) and Pronase (final concentration of 200 μg/ml, 5 hours at 37°C; Boehringer Mannheim). Unincorporated precursors and other small molecules were removed by ultrafiltration (Microcon10, 10^4-daltons cutoff; Amicon). Half of this semipurified sample was injected onto a Sephacryl S-500HR column (1 cm by 50 cm; Pharmacia) equilibrated in 0.2 M NaCl, 5 mM tris, pH 8 (0.5 ml/min, 1-ml fractions). To verify that the identity of the labeled polysaccharide was HA, we treated the other half of the original reaction with HA lyase (30 units at 37°C overnight; Sigma) before the deproteinization step. This treatment degraded the radioactive polymer to small oligosaccharides (tetramers and hexamers) that were removed by ultrafiltration before gel filtration chromatography.
28. Two cultures of NC64A cells (0.9 liter; 1.9×10^{10} cells) were infected with PBCV-1 (multiplicity of infection of 5) and incubated for 50 or 90 min after infection. Another culture served as an uninfected control. The cells were harvested, and the membrane fraction (yield ~3 mg of protein) was prepared as described [P. L. DeAngelis and A. M. Achyuthan, *J. Biol. Chem.* 271, 23657 (1996)], except that 1 mM mercaptoethanol was substituted for dithiothreitol. The paper chromatography method was used to assay for HAS activity (*13*).
29. We thank A. M. Achyuthan, G. M. Air, M. K. Brakke, R. D. Cummings, L. C. Lane, M. Nelson, and P. H. Weigel for helpful discussions. R. A. Steinberg provided the plasmids and host strain for T7 expression system. Supported by a NIH grant (R01-GM56497) and a University of Oklahoma Medical Alumni Scholarship to P.L.D. and a NIH grant (R01-GM32441) to J.V.E.

15 September 1997; accepted 30 October 1997

Specific Inhibition of Stat3 Signal Transduction by PIAS3

Chan D. Chung,* Jiayu Liao,* Bin Liu, Xiaoping Rao, Philippe Jay, Philippe Berta, Ke Shuai†

The signal transducer and activator of transcription–3 (Stat3) protein is activated by the interleukin 6 (IL-6) family of cytokines, epidermal growth factor, and leptin. A protein named PIAS3 (protein inhibitor of activated STAT) that binds to Stat3 was isolated and characterized. The association of PIAS3 with Stat3 in vivo was only observed in cells stimulated with ligands that cause the activation of Stat3. PIAS3 blocked the DNA-binding activity of Stat3 and inhibited Stat3-mediated gene activation. Although Stat1 is also phosphorylated in response to IL-6, PIAS3 did not interact with Stat1 or affect its DNA-binding or transcriptional activity. The results indicate that PIAS3 is a specific inhibitor of Stat3.

Stat3 participates in signal transduction pathways activated by the IL-6 family of cytokines and by epidermal growth factor (*1*, *2*). Stat3 is also activated in cells treated with leptin, a growth hormone that functions in regulating food intake and energy expenditure (*3*). Targeted disruption of the mouse gene encoding Stat3 leads to early embryonic lethality (*4*). Like other members of the STAT family, Stat3 becomes tyrosine phosphorylated by Janus kinases (JAKs). Phosphorylated Stat3 then forms a dimer and translocates into the nucleus to activate specific genes (*5*).

We cloned a protein named PIAS1, which can specifically interact with Stat1 (another member of the STAT family), by the yeast two-hybrid assays (*6*). We searched the expressed sequence tag (EST) database for other PIAS family members and identified a human EST clone encoding a polypeptide related to the COOH-terminal portion of PIAS1 (*7*). We obtained a full-length cDNA containing an open reading frame of 583 amino acids by screening a mouse thymus library with the human EST clone (*8*). The corresponding protein, named PIAS3, contains a putative zinc-binding motif $[C_2\text{-}(X)_{21}\text{-}C_2]$ (*9*), a feature conserved in the PIAS family (Fig. 1A). Northern (RNA) blot analysis indicated that PIAS3 is widely expressed in various human tissues (Fig. 1B).

To study the function of PIAS3, we prepared a specific antiserum (anti-PIAS3c) to a recombinant fusion protein of glutathione-S-transferase (GST) with the 79 COOH-terminal amino acid residues of PIAS3. This antibody detected a protein with a molecular mass of about 68 kD, the predicted size of PIAS3, in both cytoplasmic and nuclear extracts of a number of human and murine cell lines (*10*). To identify which STAT protein interacts with PIAS3, we prepared protein

C. D. Chung and J. Liao, Department of Biological Chemistry, University of California, Los Angeles, CA 90095, USA.
B. Liu and X. Rao, Division of Hematology-Oncology, Department of Medicine, University of California, Los Angeles, CA 90095, USA.
P. Jay and P. Berta, Centre de Recherche de Biochimie Macromoleculaire, 1919 route de Mende, 34000 Montpellier, France.
K. Shuai, Division of Hematology-Oncology, Department of Medicine; Department of Biological Chemistry; and Molecular Biology Institute, University of California, 11-934 Factor Building, 10833 LeConte Avenue, Los Angeles, CA 90095–1678, USA.
*These authors contributed equally to this manuscript.
†To whom correspondence should be addressed.

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 9/10, 15/54, 15/70, C12Q 1/68, A61K 31/715, C12P 21/00, 19/04

A3

(11) International Publication Number: WO 99/23227

(43) International Publication Date: 14 May 1999 (14.05.99)

(21) International Application Number: PCT/US98/23153

(22) International Filing Date: 30 October 1998 (30.10.98)

(30) Priority Data:

60/064,435	31 October 1997 (31.10.97)	US
09/178,851	26 October 1998 (26.10.98)	US

(71) Applicant: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA [US/US]; 1000 Asp Avenue, Norman, OK 73019 (US).

(72) Inventors: WEIGEL, Paul, H.; 817 Hollowdale, Edmond, OK 73003 (US). KUMARI, Kshama; Apartment 467, 2609 Featherstone Road, Oklahoma City, OK 73120 (US). DeANGELIS, Paul; 706 Sunnybrook Drive, Edmond, OK 73034 (US).

(74) Agents: PALMER, John et al.; Ladas & Parry, Suite 2100, 5670 Wilshire Boulevard, Los Angeles, CA 90036-5679 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
19 August 1999 (19.08.99)

(54) Title: HYALURONAN SYNTHASE GENE AND USES THEREOF

(57) Abstract

The present invention relates to a nucleic acid segment having a coding region segment encoding enzymatically active *Streptococcus equisimilis* hyaluronate synthase (seHAS), and to the use of this nucleic acid segment in the preparation of recombinant cells which produce hyaluronate synthase and its hyaluronic acid product. Hyaluronate is also known as hyaluronic acid or hyaluronan.

*(Referred to in PCT Gazette No. 4/2000, Section II)

FOR THE PURPOSES OF FORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

HYALURONAN SYNTHASE GENE AND USES THEREOF

BACKGROUND OF THE INVENTION

1. Field of the Invention.

The present invention relates to a nucleic acid segment having a coding region segment encoding enzymatically active *Streptococcus equisimilis* hyaluronate synthase (seHAS), and to the use of this nucleic acid segment in the preparation of recombinant cells which produce hyaluronate synthase and its hyaluronic acid product. Hyaluronate is also known as hyaluronic acid or hyaluronan.

2. Brief Description of the Related Art.

The incidence of streptococcal infections is a major health and economic problem worldwide, particularly in developing countries. One reason for this is due to the ability of *Streptococcal* bacteria to grow undetected by the body's phagocytic cells, i.e., macrophages and polymorphonuclear cells (PMNs). These cells are responsible for recognizing and engulfing foreign

microorganisms. One effective way ıe bacteria evade surveillance is by coating themselves with poly—accharide capsules, such as a hyaluronic acid (HA) capsule. The tructure of HA is identical in both prokaryotes and eukaryotes. ;ince HA is generally nonimmunogenic, the encapsulated b—:teria do not elicit an immune response and are, therefore, n— targeted for destruction. Moreover, the capsule exerts an a—iphagocytic effect on PMNs *in vitro* and prevents attachment of *Streptococcus* to macrophages. Precisely because of this, in Grou—A and Group C *Streptococci*, the HA capsules are major virulence fac—rs in natural and experimental infections. Group A *Streptococc—* are responsible for numerous human diseases including phary—itis, impetigo, deep tissue infections, rheumatic fever and a oxic shock-like syndrome. The Group C *Streptococcus equisimilis* i— responsible for osteomyelitis, pharyngitis, brain abscesses, and neumonia.

Structurally, HA is a h—gh molecular weight linear polysaccharide of repeating disac—haride units consisting of N-acetylglucosamine (GlcNAc) and gl—uronic acid (GlcA). The number of repeating disaccharides in an —A molecule can exceed 30,000, a $M_r>10^7$. HA is the only glycos—inogylcan synthesized by both mammalian and bacterial cells ;articularly Groups A and C *Streptococci* and Type A *Pasturell multocida*. These strains make HA which is secreted into the med—m as well as HA capsules. The mechanism by which these bacte a synthesize HA is of broad interest medicinally since the p—duction of the HA capsule is a

very efficient and clever way that *Streptococci* use to evade surveillance by the immune system.

HA is synthesized by mammalian and bacterial cells by the enzyme hyaluronate synthase which has been localized to the plasma membrane. It is believed that the synthesis of HA in these organisms is a multi-step process. Initiation involves binding of an initial precursor, UDP-GlcNAc or UDP-GlcA. This is followed by elongation which involves alternate addition of the two sugars to the growing oligosaccharide chain. The growing polymer is extruded across the plasma membrane region of the cell and into the extracellular space. Although the HA biosynthetic system was one of the first membrane heteropolysaccharide synthetic pathways studied, the mechanism of HA synthesis is still not well understood. This may be because *in vitro* systems developed to date are inadequate in that *de novo* biosynthesis of HA has not been accomplished.

The direction of HA polymer growth is still a matter of disagreement among those of ordinary skill in the art. Addition of the monosaccharides could be to the reducing or nonreducing end of the growing HA chain. Furthermore, questions remain concerning (i) whether nascent chains are linked covalently to a protein, to UDP or to a lipid intermediate, (ii) whether chains are initiated using a primer, and (iii) the mechanism by which the mature polymer is extruded through the plasma membrane of the *Streptococcus*. Understanding the mechanism of HA biosynthesis may allow

development of alternative strategies to control *Streptococcal* and *Pasturella* infections by interfering in the process.

HA has been identified in virtually every tissue in vertebrates and has achieved widespread use in various clinical applications, most notably and appropriately as an intra-articular matrix supplement and in eye surgery. The scientific literature has also shown a transition from the original perception that HA is primarily a passive structural component in the matrix of a few connective tissues and in the capsule of certain strains of bacteria to a recognition that this ubiquitous macromolecule is dynamically involved in many biological processes: from modulating cell migration and differentiation during embryogenesis to regulation of extracellular matrix organization and metabolism to important roles in the complex processes of metastasis, wound healing, and inflammation. Furthermore it is becoming clear that HA is highly metabolically active and that cells focus much attention on the processes of its synthesis and catabolism. For example, the half-life of HA in tissues ranges from 1 to 3 weeks in cartilage to <1 day in epidermis.

It is now clear that a single protein utilizes both sugar substrates to synthesize HA. The abbreviation HAS, for the HA synthase, has gained widespread support for designating this class of enzymes. Markovitz *et al.* successfully characterized the HAS activity from *Streptococcus pyogenes* and discovered the enzymes's membrane localization and its requirements for sugar nucleotide precursors and Mg^{2+}. Prehm found that elongating HA, made by B6

cells, was digested by hyaluronidase added to the medium and proposed that HAS resides at the plasma membrane. Philipson and Schwartz also showed that HAS activity cofractionated with plasma membrane markers in mouse oligodendroglioma cells.

HAS assembles high M_r HA that is simultaneously extruded through the membrane into the extracellular space (or to make the cell capsule in the case of bacteria) as glycosaminoglycan synthesis proceeds. This mode of biosynthesis is unique among macromolecules since nucleic acids, proteins, and lipids are synthesized in the nucleus, endoplasmic reticulum/Golgi, cytoplasm, or mitochondria. The extrusion of the growing chain into the extracellular space also allows for unconstrained polymer growth, thereby achieving the exceptionally large size of HA, whereas confinement of synthesis within a Golgi or post-Golgi compartment could limit the overall amount or length of the polymers formed. High concentrations of HA within a confined lumen could also create a high viscosity environment that might be deleterious for other organelle functions.

Several studies attempted to solubilize, identify, and purify HAS from strains of *Streptococci* that make a capsular coat of HA as well as from eukaryotic cells. Although the streptococcal and murine oligodendroglioma enzymes were successfully detergent-solubilized and studied, efforts to purify an active HAS for further study or molecular cloning remained unsuccessful for decades. Prehm and Mausolf used periodate-oxidized UDP-GlcA or UDP-GlcNAc to affinity label a protein of ~52 kDa in streptococcal

membranes that co-purified with HAS This led to a report claiming that the Group C streptococcal HA had been cloned, which was unfortunately erroneous. This tudy failed to demonstrate expression of an active synthase and may have actually cloned a peptide transporter. Triscott and van de Rijn used digitonin to solubilize HAS from streptococcal m branes in an active form. Van de Rijn and Drake selectively ra olabeled three streptococcal membrane proteins of 42, 33, and 2 kDa with 5-azido-UDP-GlcA and suggested that the 33-kDa protei was HAS. As shown later, however, HAS actually turned out t be the 42-kDa protein.

Despite these efforts, pr ress in understanding the regulation and mechanisms of HA sy nesis was essentially stalled, since there were no molecular prob s for HAS mRNA or HAS protein. A major breakthrough occurred i 1993 when DeAngelis *et al.* reported the molecular cloning and haracterization of the Group A streptococcal gene encoding the pr ein HasA. This gene was known to be in part of an operon requi d for bacterial HA synthesis, although the function of this pro in, which is now designated as spHAS (the S. *pyogenes* HAS), was known. spHAS was subsequently proven to be responsible for HA elongation and was the first glycosaminoglycan synthase iden ified and cloned and then successfully expressed. The S. *pyogenes* HA synthesis operon encodes two other proteins. HasB s a UDP-glucose dehydrogenase, which is required to convert UDP ucose to UDP-GlcA, one of the substrates for HA synthesis. HasC is a UDP-glucose pyrophosphorylase, which is requir to convert glucose 1-phosphate

and UTP to UDP-glucose. Co-transfection of both *hasA* and *hasB* genes into either acapsular *Streptococcus* strains or *Enteroccus faecalis* conferred them with the ability to synthesize HA and form a capsule. This provided the first strong evidence that HasA is an HA synthase.

The elusive HA synthase gene was finally cloned by a transposon mutagenesis approach, in which an acapsular mutant Group A strain was created containing a transposon interruption of the HA synthesis operon. Known sequences of the transposon allowed the region of the junction with streptococcal DNA to be identified and then cloned from wild-type cells. The encoded spHAS was 5-10% identical to a family of yeast chitin synthases and 30% identical to the *Xenopus laevis* protein DG42 (developmentally expressed during gastrulation), whose function was unknown at the time. DeAngelis and Weigel expressed the active recombinant spHAS in *Escherichia coli* and showed that this single purified gene product synthesizes high M_r HA when incubated in vitro with UDP-GlcA and UDP-GlcNAc, thereby showing that both glycosyltransferase activities required for HA synthesis are catalyzed by the same protein, as first proposed in 1959. This set the stage for the almost simultaneous identification of cukaryotic HAS cDNAs in 1996 by four laboratories revealing that HAS is a multigene family encoding distinct isozymes. Two genes (*HAS1* and *HAS2*) were quickly discovered in mammals (29-34), and a third gene *HAS3* was later discovered. A second streptococcal seHAS or *Streptococcus*

equisimilis hyaluronate synthase, as now been found and is the invention being claimed and disclo :d herein.

As indicated, we have also id ıtified the authentic HAS gene from Group C *Streptococcus equisim is* (seHAS); the seHAS protein has a high level of identity (ap oximately 70 percent) to the spHAS enzyme. This identity, howe r, is interesting because the seHAS gene does not cross-hybridiz to the spHAS gene.

Membranes prepared from *E. c* i expressing recombinant seHAS synthesize HA when both substrat : are provided. The results confirm that the earlier report of ansing *et al.* claiming to have cloned the Group C HAS was wrong. Jnfortunately, several studies have employed antibody to this unch acterized 52-kDa streptococcal protein to investigate what was b ieved to be eukaryotic HAS.

Itano and Kimata used expre ion cloning in a mutant mouse mammary carcinoma cell line, unabl to synthesize HA, to clone the first putative mammalian HAS cDNA mHAS1). Subclones defective in HA synthesis fell into three separate classes that were complementary for HA synthesis in ›matic cell fusion experiments, suggesting that at least three pro ins are required. Two of these classes maintained some HA synthe .c activity, whereas one showed none. The latter cell line was ısed in transient transfection experiments with cDNA prepared fr the parental cells to identify a single protein that restored H synthetic activity. Sequence analyses revealed a deduced prima structure for a protein of ~65 kDa with a predicted membrane to ılogy similar to that of spHAS. mmHAS1 is 30% identical to spHAS ınd 55% identical to DG42. The

same month this report appeared, three other groups submitted papers describing cDNAs encoding what was initially thought to be the same mouse and human enzyme. However, through an extraordinary circumstance, each of the four laboratories had discovered a separate HAS isozyme in both species.

Using a similar functional cloning approach to that of Itano and Kimata, Shyjan *et al.* identified the human homolog of HAS 1. A mesenteric lymph node cDNA library was used to transfect murine mucosal T lymphocytes that were then screened for their ability to adhere in a rosette assay. Adhesion of one transfectant was inhibited by antisera to CD44, a known cell surface HA-binding protein, and was abrogated directly by pretreatment with hyaluronidase. Thus, rosetting by this transfectant required synthesis of HA. Cloning and sequencing of the responsible cDNA identified hsHAS1. Itano and Kimata also reported a human HAS1 cDNA isolated from a fetal brain library. The hsHAS1 cDNAs reported by the two groups, however, differ in length; they encode a 578 or a 543 amino acid protein. HAS activity has only been demonstrated for the longer form.

Based on the molecular identification of spHAS as an authentic HA synthase and regions of near identity among DG42, spHAS, and NodC (a β-GlcNAc transferase nodulation factor in *Rhizobium*), Spicer *et al.* used a degenerate RT-PCR approach to clone a mouse embryo cDNA encoding a second distinct enzyme, which is designated mmHAS2. Transfection of mmHAS2 cDNA into COS cells directed *de novo* production of an HA cell coat detected by a particle exclusion

assay, thereby providing strong evi_ence that the HAS2 protein can synthesize HA. Using a similar ap=roach, Watanabe and Yamaguchi screened a human fetal brain cDN_ library to identify hsHAS2. Fulop *et al.* independently used =similar strategy to identify mmHAS2 in RNA isolated from ov=ian cumulus cells actively synthesizing HA, a critical proce= for normal cumulus oophorus expansion in the pre-ovulatory f=licle. Cumulus cell-oocyte complexes were isolated from mice mediately after initiating an ovulatory cycle, before HA synthe s begins, and at later times when HA synthesis is just beginnir= (3 h) or already apparent (4 h). RT-PCR showed that HAS2 m=A was absent initially but expressed at high levels 3-4 h late= suggesting that transcription of HAS2 regulates HA synthesis in his process. Both hsHAS2 are 552 amino acids in length and ar=98% identical. mmHAS1 is 583 amino acids long an 95% identical to hsHAS1, which is 578 amino acids long.

Most recently Spicer *et al.* u=d a PCR approach to identify a third HAS gene in mammals. The m=AS3 protein is 554 amino acids long and 71, 56, and 28% ident al, respectively, to mmHAS1, mmHAS2, DG42, and spHAS. Spicer t *al.* have also localized the three human and mouse genes to thr= different chromosomes (HAS1 to hsChr 19/mmChr 17; HAS2 to hsChr 8 mChr 15; HAS3 to hsChr 16/mmChr 8). Localization of the three HA genes on different chromosomes and the appearance of HA throughout the vertebrate class suggest that this gene family is ancient and that isozymes appeared by duplication early in the evolution of vertebrates. The high

identity (~30%) between the bacterial and eukaryotic HASs also suggests that the two had a common ancestral gene. Perhaps primitive bacteria usurped the HAS gene from an early vertebrate ancestor before the eukaryotic gene products became larger and more complex. Alternatively, the bacteria could have obtained a larger vertebrate HAS gene and deleted regulatory sequences nonessential for enzyme activity.

The discovery of X. *laevis* DG42 by Dawid and co-workers played a significant role in these recent developments, even though this protein was not known to be an HA synthase. Nonetheless, that DG42 and spHAS were 30% identical was critical for designing oligonucleotides that allowed identification of mammalian HAS2. Ironically, definitive evidence that DG42 is a *bona fide* HA synthase was reported only after the discoveries of the Mammalian isozymes, when DeAngelis and Achyuthan expressed the recombinant protein in yeast (an organism that cannot synthesize HA) and showed that it synthesizes HA when isolated membranes are provided with the two substrates. Meyer and Kreil also showed that lysates from cells transfected with cDNA for DG42 synthesize elevated levels of HA. Now that its function is known, DG42 can, therefore, be designated XlHAS.

There are common predicted structural features shared by all the HAS proteins, including a large central domain and clusters of 2-3 transmembrane or membrane-associated domains at both the amino and carboxyl ends of the protein. The central domain, which comprises up to ~88% of the predicted intracellular HAS protein

sequences, probably contains the catalytic regions of the enzyme. This predicted central domain is 26[illegible] amino acids long in spHAS (63% of the total protein) and 307-328 residues long in the eukaryotic HAS members (54-56% of the total protein). The exact number and orientation of membrane domains and the topological organization of extracellular and intracellular loops have not yet been experimentally determined for any HAS.

spHAS is a HAS family member that has been purified and partially characterized. Initial studies using spHAS/alkaline phosphatase fusion proteins indicate that the N terminus, C terminus, and the large central domain of spHAS are, in fact, inside the cell. spHAS has 6 cysteines, whereas HAS1, HAS2, and HAS3 have 13, 14 and 14 Cys residues, respectively. Two of the 6 Cys residues in spHAS are conserved and identical in HAS1 and HAS2. Only one conserved Cys residue is found at the same position (Cys-225 in spHAS) in all the HAS family members. This may be an essential Cys whose modification by sulfhydryl poisons partially inhibits enzyme activity. The possible presence of disulfide bonds or the identification of critical Cys residues needed for any of the multiple HAS functions noted below has not yet been elucidated for any members of the HAS family.

In addition to the proposed unique mode of synthesis at the plasma membrane, the HAS enzyme family is highly unusual in the large number of functions required for the overall polymerization of HA. At least six discrete activities are present within the HAS enzyme: binding sites for each of the two different sugar

nucleotide precursors (UDP-GlcNAc and UDP-GlcA), two different glycosyltransferase activities, one or more binding sites that anchor the growing HA polymer to the enzyme (perhaps related to a B-X_7-B motif), and a ratchet-like transfer reaction that moves the growing polymer one sugar at a time. This later activity is likely coincident with the stepwise advance of the polymer through the membrane. All of these functions, and perhaps others as yet unknown, are present in a relatively small protein ranging in size from 419 (spHAS) to 588 (xHAS) amino acids.

Although all the available evidence supports the conclusion that only the spHAS protein is required for HA biosynthesis in bacteria or *in vitro*, it is possible that the larger eukaryotic HAS family members are part of multicomponent complexes. Since the eukaryotic HAS proteins are ~40% larger than spHAS, their additional protein domains could be involved in more elaborate functions such as intracellular trafficking and localization, regulation of enzyme activity, and mediating interactions with other cellular components.

The unexpected finding that there are multiple vertebrate HAS genes encoding different synthases strongly supports the emerging consensus that HA is an important regulator of cell behavior and not simply a structural component in tissues. Thus, in less than six months, the field moved from one known, cloned HAS (spHAS) to recognition of a multigene family that promises rapid, numerous, and exciting future advances in our understanding of the synthesis and biology of HA.

For example, disclosed herei ıfter are the sequences of the two HAS genes: from *Pasturella ultocida;* and (2) *Paramecium bursaria chlorella virus* (PBCV-1 The presence of hyaluronan synthase in these two systems and he purification and use of the hyaluronan synthase from these tw=different systems indicates an ability to purify and isolate r=leic acid sequences encoding enzymatically active hyaluronan= synthase in many different prokaryotic and viral sources.

Group C *Streptococcus equisi=lis* strain D181 synthesizes and secretes hyaluronic acid (HA). In—stigators have used this strain and Group A *Streptococcus pyogene* :rains, such as S43 and A111, to study the biosynthesis of HA ınd to characterize the HA-synthesizing activity in terms of :s divalent cation requirement, precursor (UDP-GlcNAc and UDP-Glc utilization, and optimum pH.

Traditionally, HA has been p=pared commercially by isolation from either rooster combs or extra llular media from Streptococcal cultures. One method which has b=n developed for preparing HA is through the use of cultures of HA-=oducing Streptococcal bacteria. U.S. Patent No. 4,517,295 descri■s such a procedure wherein HA-producing Streptococci are fermen :d under anaerobic conditions in a CO_2-enriched growth medium. Inder these conditions, HA is produced and can be extracted fro=the broth. It is generally felt that isolation of HA from rooster :mbs is laborious and difficult, since one starts with HA in a l=s pure state. The advantage of isolation from rooster combs is ■at the HA produced is of higher molecular weight. However, p=paration of HA by bacterial

fermentation is easier, since the HA is of higher purity to start with. Usually, however, the molecular weight of HA produced in this way is smaller than that from rooster combs. Therefore, a technique that would allow the production of high molecular weight HA by bacterial fermentation would be an improvement over existing procedures.

High molecular weight HA has a wide variety of useful applications -- ranging from cosmetics to eye surgery. Due to its potential for high viscosity and its high biocompatibility, HA finds particular application in eye surgery as a replacement for vitreous fluid. HA has also been used to treat racehorses for traumatic arthritis by intra-articular injections of HA, in shaving cream as a lubricant, and in a variety of cosmetic products due to its physiochemical properties of high viscosity and its ability to retain moisture for long periods of time. In fact, in August of 1997 the U.S. Food and Drug Agency approved the use of high molecular weight HA in the treatment of severe arthritis through the injection of such high molecular weight HA directly into the affected joints. In general, the higher molecular weight HA that is employed the better. This is because HA solution viscosity increases with the average molecular weight of the individual HA polymer molecules in the solution. Unfortunately, very high molecular weight HA, such as that ranging up to 10^7, has been difficult to obtain by currently available isolation procedures.

To address these or other difficulties, there is a need for new methods and constructs that can be used to produce HA having

one or more improved properties such as greater purity or ease of preparation. In particular, there is a need to develop methodology for the production of larger amounts of relatively high molecular weight and relatively pure HA than is currently commercially available. There is yet another need to be able to develop methodology for the production of HA having a modified size distribution ($HA_{\Delta size}$) as well as HA having a modified structure ($HA_{\Delta mod}$).

The present invention addresses one or more shortcomings in the art. Using recombinant DNA technology, a purified nucleic acid segment having a coding region encoding enzymatically active seHAS is disclosed and claimed in conjunction, with methods to produce an enzymatically active HA synthase, as well as methods for using the nucleic acid segment in the preparation of recombinant cells which produce HAS and its hyaluronic acid product.

Thus, it is an object of the present invention to provide a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

It is a further object of the present invention to provide a recombinant vector which includes a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

It is still a further object of the present invention to provide a recombinant host cell transformed with a recombinant vector which includes a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

It is yet another object of the present invention to provide a method for detecting a bacterial cell that expresses HAS.

It is another object of the present invention to provide a method for producing high and/or low molecular weight hyaluronic acid from a hyaluronate synthase gene, such as seHAS, as well as methods for producing HA having a modified size distribution and/or a modified structure.

These and other objects of the present invention will become apparent in light of the attached specification, claims, and drawings.

BRIEF SUMMARY OF THE INVENTION

The present invention involves the application of recombinant DNA technology to solving one or more problems in the art of hyaluronic acid (HA) preparation. These problems are addressed through the isolation and use of a nucleic acid segment having a coding region encoding the enzymatically active *Streptococcus equisimilis* (seHAS) hyaluronate synthase gene, a gene responsible for HA chain biosynthesis. The seHAS gene was cloned from DNA of an appropriate microbial source and engineered into useful recombinant constructs for the preparation of HA and for the preparation of large quantities of the HAS enzyme itself.

The present invention encompasses a novel gene, seHAS. The expression of this gene correlates with virulence of *Streptococcal* Group A and Group C strains, by providing a means of escaping phagocytosis and immune surveillance. The terms "hyaluronic acid synthase", "hyaluronate synthase", "hyaluronan synthase" and "HA

synthase", are used interchangeably to describe an enzyme that polymerizes a glycosaminoglycan polysaccharide chain composed of alternating glucuronic acid and N-acetylglucosamine sugars, β 1,3 and β 1,4 linked. The term "seHAS" describes the HAS enzyme derived from *Streptococcus equisimilis*.

The present invention concerns the isolation and characterization of a hyaluronate or hyaluronic acid synthase gene, cDNA, and gene product (HAS), as may be used for the polymerization of glucuronic acid and N-acetylglucosamine into the glycosaminoglycan hyaluronic acid. The present invention identifies the seHAS locus and discloses the nucleic acid sequence which encodes for the enzymatically active seHAS gene from *Streptococcus equisimilis*. The HAS gene also provides a new probe to assess the potential of bacterial specimens to produce hyaluronic acid.

Through the application of techniques and knowledge set forth herein, those of skill in the art will be able to obtain nucleic acid segments encoding the seHAS gene. As those of skill in the art will recognize, in light of the present disclosure, these advantages provide significant utility in being able to control the expression of the seHAS gene and control the nature of the seHAS gene product, the seHAS enzyme, that is produced.

Accordingly, the invention is directed to the isolation of a purified nucleic acid segment which has a coding region encoding enzymatically active HAS, whether it be from prokaryotic or eukaryotic sources. This is possible because the enzyme, and

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 9/10, 15/54, 15/70, C12Q 1/68, A61K 31/715, C12P 21/00, 19/04

A3

(11) International Publication Number: WO 99/23227

(43) International Publication Date: 14 May 1999 (14.05.99)

(21) International Application Number: PCT/US98/23153

(22) International Filing Date: 30 October 1998 (30.10.98)

(30) Priority Data:

60/064,435	31 October 1997 (31.10.97)	US
09/178,851	26 October 1998 (26.10.98)	US

(71) Applicant: THE BOARD OF REGENTS OF THE UNIVERSITY OF OKLAHOMA [US/US]; 1000 Asp Avenue, Norman, OK 73019 (US).

(72) Inventors: WEIGEL, Paul, H.; 817 Hollowdale, Edmond, OK 73003 (US). KUMARI, Kshama; Apartment 467, 2609 Featherstone Road, Oklahoma City, OK 73120 (US). DeANGELIS, Paul; 706 Sunnybrook Drive, Edmond, OK 73034 (US).

(74) Agents: PALMER, John et al.; Ladas & Parry, Suite 2100, 5670 Wilshire Boulevard, Los Angeles, CA 90036–5679 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
19 August 1999 (19.08.99)

(54) Title: HYALURONAN SYNTHASE GENE AND USES THEREOF

(57) Abstract

The present invention relates to a nucleic acid segment having a coding region segment encoding enzymatically active *Streptococcus equisimilis* hyaluronate synthase (seHAS), and to the use of this nucleic acid segment in the preparation of recombinant cells which produce hyaluronate synthase and its hyaluronic acid product. Hyaluronate is also known as hyaluronic acid or hyaluronan.

*(Referred to in PCT Gazette No. 4/2000, Section II)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

HYALURONAN SYNTHASE GENE AND USES THEREOF

BACKGROUND OF THE INVENTION

1. Field of the Invention.

The present invention relates to a nucleic acid segment having a coding region segment encoding enzymatically active *Streptococcus equisimilis* hyaluronate synthase (seHAS), and to the use of this nucleic acid segment in the preparation of recombinant cells which produce hyaluronate synthase and its hyaluronic acid product. Hyaluronate is also known as hyaluronic acid or hyaluronan.

2. Brief Description of the Related Art.

The incidence of streptococcal infections is a major health and economic problem worldwide, particularly in developing countries. One reason for this is due to the ability of *Streptococcal* bacteria to grow undetected by the body's phagocytic cells, i.e., macrophages and polymorphonuclear cells (PMNs). These cells are responsible for recognizing and engulfing foreign

microorganisms. One effective way ıe bacteria evade surveillance is by coating themselves with poly—accharide capsules, such as a hyaluronic acid (HA) capsule. The tructure of HA is identical in both prokaryotes and eukaryotes. ;ince HA is generally nonimmunogenic, the encapsulated b :teria do not elicit an immune response and are, therefore, n— targeted for destruction. Moreover, the capsule exerts an a—iphagocytic effect on PMNs *in vitro* and prevents attachment of *Streptococcus* to macrophages. Precisely because of this, in Grou— and Group C *Streptococci*, the HA capsules are major virulence fac—rs in natural and experimental infections. Group A *Streptococc*— are responsible for numerous human diseases including phary—itis, impetigo, deep tissue infections, rheumatic fever and a oxic shock-like syndrome. The Group C *Streptococcus equisimilis* i— responsible for osteomyelitis, pharyngitis, brain abscesses, and neumonia.

Structurally, HA is a h—gh molecular weight linear polysaccharide of repeating disac—haride units consisting of N-acetylglucosamine (GlcNAc) and gl—uronic acid (GlcA). The number of repeating disaccharides in an — molecule can exceed 30,000, a $M_r>10^7$. HA is the only glycos—inogylcan synthesized by both mammalian and bacterial cells articularly Groups A and C *Streptococci* and Type A *Pasturell multocida*. These strains make HA which is secreted into the med—m as well as HA capsules. The mechanism by which these bacte a synthesize HA is of broad interest medicinally since the p—duction of the HA capsule is a

very efficient and clever way that *Streptococci* use to evade surveillance by the immune system.

HA is synthesized by mammalian and bacterial cells by the enzyme hyaluronate synthase which has been localized to the plasma membrane. It is believed that the synthesis of HA in these organisms is a multi-step process. Initiation involves binding of an initial precursor, UDP-GlcNAc or UDP-GlcA. This is followed by elongation which involves alternate addition of the two sugars to the growing oligosaccharide chain. The growing polymer is extruded across the plasma membrane region of the cell and into the extracellular space. Although the HA biosynthetic system was one of the first membrane heteropolysaccharide synthetic pathways studied, the mechanism of HA synthesis is still not well understood. This may be because *in vitro* systems developed to date are inadequate in that *de novo* biosynthesis of HA has not been accomplished.

The direction of HA polymer growth is still a matter of disagreement among those of ordinary skill in the art. Addition of the monosaccharides could be to the reducing or nonreducing end of the growing HA chain. Furthermore, questions remain concerning (i) whether nascent chains are linked covalently to a protein, to UDP or to a lipid intermediate, (ii) whether chains are initiated using a primer, and (iii) the mechanism by which the mature polymer is extruded through the plasma membrane of the *Streptococcus*. Understanding the mechanism of HA biosynthesis may allow

development of alternative strategies to control *Streptococcal* and *Pasturella* infections by interfering in the process.

HA has been identified in virtually every tissue in vertebrates and has achieved widespread use in various clinical applications, most notably and appropriately as an intra-articular matrix supplement and in eye surgery. The scientific literature has also shown a transition from the original perception that HA is primarily a passive structural component in the matrix of a few connective tissues and in the capsule of certain strains of bacteria to a recognition that this ubiquitous macromolecule is dynamically involved in many biological processes: from modulating cell migration and differentiation during embryogenesis to regulation of extracellular matrix organization and metabolism to important roles in the complex processes of metastasis, wound healing, and inflammation. Furthermore it is becoming clear that HA is highly metabolically active and that cells focus much attention on the processes of its synthesis and catabolism. For example, the half-life of HA in tissues ranges from 1 to 3 weeks in cartilage to <1 day in epidermis.

It is now clear that a single protein utilizes both sugar substrates to synthesize HA. The abbreviation HAS, for the HA synthase, has gained widespread support for designating this class of enzymes. Markovitz *et al.* successfully characterized the HAS activity from *Streptococcus pyogenes* and discovered the enzymes's membrane localization and its requirements for sugar nucleotide precursors and Mg^{2+}. Prehm found that elongating HA, made by B6

cells, was digested by hyaluronidase added to the medium and proposed that HAS resides at the plasma membrane. Philipson and Schwartz also showed that HAS activity cofractionated with plasma membrane markers in mouse oligodendroglioma cells.

HAS assembles high M_r HA that is simultaneously extruded through the membrane into the extracellular space (or to make the cell capsule in the case of bacteria) as glycosaminoglycan synthesis proceeds. This mode of biosynthesis is unique among macromolecules since nucleic acids, proteins, and lipids are synthesized in the nucleus, endoplasmic reticulum/Golgi, cytoplasm, or mitochondria. The extrusion of the growing chain into the extracellular space also allows for unconstrained polymer growth, thereby achieving the exceptionally large size of HA, whereas confinement of synthesis within a Golgi or post-Golgi compartment could limit the overall amount or length of the polymers formed. High concentrations of HA within a confined lumen could also create a high viscosity environment that might be deleterious for other organelle functions.

Several studies attempted to solubilize, identify, and purify HAS from strains of *Streptococci* that make a capsular coat of HA as well as from eukaryotic cells. Although the streptococcal and murine oligodendroglioma enzymes were successfully detergent-solubilized and studied, efforts to purify an active HAS for further study or molecular cloning remained unsuccessful for decades. Prehm and Mausolf used periodate-oxidized UDP-GlcA or UDP-GlcNAc to affinity label a protein of ~52 kDa in streptococcal

membranes that co-purified with HAS This led to a report claiming that the Group C streptococcal HA had been cloned, which was unfortunately erroneous. This tudy failed to demonstrate expression of an active synthase and may have actually cloned a peptide transporter. Triscott and van de Rijn used digitonin to solubilize HAS from streptococcal membranes in an active form. Van de Rijn and Drake selectively radiolabeled three streptococcal membrane proteins of 42, 33, and 2 kDa with 5-azido-UDP-GlcA and suggested that the 33-kDa protein was HAS. As shown later, however, HAS actually turned out to be the 42-kDa protein.

Despite these efforts, progress in understanding the regulation and mechanisms of HA synthesis was essentially stalled, since there were no molecular probes for HAS mRNA or HAS protein. A major breakthrough occurred in 1993 when DeAngelis *et al.* reported the molecular cloning and characterization of the Group A streptococcal gene encoding the protein HasA. This gene was known to be in part of an operon required for bacterial HA synthesis, although the function of this protein, which is now designated as spHAS (the S. *pyogenes* HAS), was unknown. spHAS was subsequently proven to be responsible for HA elongation and was the first glycosaminoglycan synthase identified and cloned and then successfully expressed. The S. *pyogenes* HA synthesis operon encodes two other proteins. HasB is a UDP-glucose dehydrogenase, which is required to convert UDP-glucose to UDP-GlcA, one of the substrates for HA synthesis. HasC is a UDP-glucose pyrophosphorylase, which is required to convert glucose 1-phosphate

and UTP to UDP-glucose. Co-transfection of both *hasA* and *hasB* genes into either acapsular *Streptococcus* strains or *Enteroccus faecalis* conferred them with the ability to synthesize HA and form a capsule. This provided the first strong evidence that HasA is an HA synthase.

The elusive HA synthase gene was finally cloned by a transposon mutagenesis approach, in which an acapsular mutant Group A strain was created containing a transposon interruption of the HA synthesis operon. Known sequences of the transposon allowed the region of the junction with streptococcal DNA to be identified and then cloned from wild-type cells. The encoded spHAS was 5-10% identical to a family of yeast chitin synthases and 30% identical to the *Xenopus laevis* protein DG42 (developmentally expressed during gastrulation), whose function was unknown at the time. DeAngelis and Weigel expressed the active recombinant spHAS in *Escherichia coli* and showed that this single purified gene product synthesizes high M_r HA when incubated *in vitro* with UDP-GlcA and UDP-GlcNAc, thereby showing that both glycosyltransferase activities required for HA synthesis are catalyzed by the same protein, as first proposed in 1959. This set the stage for the almost simultaneous identification of eukaryotic HAS cDNAs in 1996 by four laboratories revealing that HAS is a multigene family encoding distinct isozymes. Two genes (*HAS1* and *HAS2*) were quickly discovered in mammals (29-34), and a third gene *HAS3* was later discovered. A second streptococcal seHAS or *Streptococcus*

equisimilis hyaluronate synthase, as now been found and is the invention being claimed and disclo d herein.

As indicated, we have also id tified the authentic HAS gene from Group C *Streptococcus equisim is* (seHAS); the seHAS protein has a high level of identity (ap oximately 70 percent) to the spHAS enzyme. This identity, howe r, is interesting because the seHAS gene does not cross-hybridiz to the spHAS gene.

Membranes prepared from *E. c* i expressing recombinant seHAS synthesize HA when both substrat s are provided. The results confirm that the earlier report of ansing *et al.* claiming to have cloned the Group C HAS was wrong. Unfortunately, several studies have employed antibody to this unch acterized 52-kDa streptococcal protein to investigate what was b ieved to be eukaryotic HAS.

Itano and Kimata used expre ion cloning in a mutant mouse mammary carcinoma cell line, unabl to synthesize HA, to clone the first putative mammalian HAS cDNA mHAS1). Subclones defective in HA synthesis fell into three separate classes that were complementary for HA synthesis in matic cell fusion experiments, suggesting that at least three pro ins are required. Two of these classes maintained some HA synthe c activity, whereas one showed none. The latter cell line was sed in transient transfection experiments with cDNA prepared fr the parental cells to identify a single protein that restored H synthetic activity. Sequence analyses revealed a deduced prima structure for a protein of ~65 kDa with a predicted membrane to logy similar to that of spHAS. mmHAS1 is 30% identical to spHAS nd 55% identical to DG42. The

same month this report appeared, three other groups submitted papers describing cDNAs encoding what was initially thought to be the same mouse and human enzyme. However, through an extraordinary circumstance, each of the four laboratories had discovered a separate HAS isozyme in both species.

Using a similar functional cloning approach to that of Itano and Kimata, Shyjan *et al.* identified the human homolog of HAS 1. A mesenteric lymph node cDNA library was used to transfect murine mucosal T lymphocytes that were then screened for their ability to adhere in a rosette assay. Adhesion of one transfectant was inhibited by antisera to CD44, a known cell surface HA-binding protein, and was abrogated directly by pretreatment with hyaluronidase. Thus, rosetting by this transfectant required synthesis of HA. Cloning and sequencing of the responsible cDNA identified hsHAS1. Itano and Kimata also reported a human HAS1 cDNA isolated from a fetal brain library. The hsHAS1 cDNAs reported by the two groups, however, differ in length; they encode a 578 or a 543 amino acid protein. HAS activity has only been demonstrated for the longer form.

Based on the molecular identification of spHAS as an authentic HA synthase and regions of near identity among DG42, spHAS, and NodC (a β-GlcNAc transferase nodulation factor in *Rhizobium*), Spicer *et al.* used a degenerate RT-PCR approach to clone a mouse embryo cDNA encoding a second distinct enzyme, which is designated mmHAS2. Transfection of mmHAS2 cDNA into COS cells directed *de novo* production of an HA cell coat detected by a particle exclusion

assay, thereby providing strong evi__nce that the HAS2 protein can synthesize HA. Using a similar ap__roach, Watanabe and Yamaguchi screened a human fetal brain cDN__ library to identify hsHAS2. Fulop *et al.* independently used __similar strategy to identify mmHAS2 in RNA isolated from ov__ian cumulus cells actively synthesizing HA, a critical proce__ for normal cumulus oophorus expansion in the pre-ovulatory f__licle. Cumulus cell-oocyte complexes were isolated from mice __mediately after initiating an ovulatory cycle, before HA synthe__s begins, and at later times when HA synthesis is just beginnir__ (3 h) or already apparent (4 h). RT-PCR showed that HAS2 m__NA was absent initially but expressed at high levels 3-4 h late__ suggesting that transcription of HAS2 regulates HA synthesis in his process. Both hsHAS2 are 552 amino acids in length and are__98% identical. mmHAS1 is 583 amino acids long an 95% identical__to hsHAS1, which is 578 amino acids long.

Most recently Spicer *et al.* u__d a PCR approach to identify a third HAS gene in mammals. The m__AS3 protein is 554 amino acids long and 71, 56, and 28% ident__al, respectively, to mmHAS1, mmHAS2, DG42, and spHAS. Spicer __t al. have also localized the three human and mouse genes to thre__ different chromosomes (HAS1 to hsChr 19/mmChr 17; HAS2 to hsChr 8__mChr 15; HAS3 to hsChr 16/mmChr 8). Localization of the three HA__genes on different chromosomes and the appearance of HA throughc__t the vertebrate class suggest that this gene family is ancien__and that isozymes appeared by duplication early in the evolut__on of vertebrates. The high

identity (~30%) between the bacterial and eukaryotic HASs also suggests that the two had a common ancestral gene. Perhaps primitive bacteria usurped the HAS gene from an early vertebrate ancestor before the eukaryotic gene products became larger and more complex. Alternatively, the bacteria could have obtained a larger vertebrate HAS gene and deleted regulatory sequences nonessential for enzyme activity.

The discovery of X. *laevis* DG42 by Dawid and co-workers played a significant role in these recent developments, even though this protein was not known to be an HA synthase. Nonetheless, that DG42 and spHAS were 30% identical was critical for designing oligonucleotides that allowed identification of mammalian HAS2. Ironically, definitive evidence that DG42 is a *bona fide* HA synthase was reported only after the discoveries of the Mammalian isozymes, when DeAngelis and Achyuthan expressed the recombinant protein in yeast (an organism that cannot synthesize HA) and showed that it synthesizes HA when isolated membranes are provided with the two substrates. Meyer and Kreil also showed that lysates from cells transfected with cDNA for DG42 synthesize elevated levels of HA. Now that its function is known, DG42 can, therefore, be designated XlHAS.

There are common predicted structural features shared by all the HAS proteins, including a large central domain and clusters of 2-3 transmembrane or membrane-associated domains at both the amino and carboxyl ends of the protein. The central domain, which comprises up to ~88% of the predicted intracellular HAS protein

sequences, probably contains the catalytic regions of the enzyme. This predicted central domain is 26[illegible] amino acids long in spHAS (63% of the total protein) and 307-328 residues long in the eukaryotic HAS members (54-56% of the total protein). The exact number and orientation of membrane domains and the topological organization of extracellular and intracellular loops have not yet been experimentally determined for any HAS.

spHAS is a HAS family member that has been purified and partially characterized. Initial studies using spHAS/alkaline phosphatase fusion proteins indicate that the N terminus, C terminus, and the large central domain of spHAS are, in fact, inside the cell. spHAS has 6 cysteines, whereas HAS1, HAS2, and HAS3 have 13, 14 and 14 Cys residues, respectively. Two of the 6 Cys residues in spHAS are conserved and identical in HAS1 and HAS2. Only one conserved Cys residue is found at the same position (Cys-225 in spHAS) in all the HAS family members. This may be an essential Cys whose modification by sulfhydryl poisons partially inhibits enzyme activity. The possible presence of disulfide bonds or the identification of critical Cys residues needed for any of the multiple HAS functions noted below has not yet been elucidated for any members of the HAS family.

In addition to the proposed unique mode of synthesis at the plasma membrane, the HAS enzyme family is highly unusual in the large number of functions required for the overall polymerization of HA. At least six discrete activities are present within the HAS enzyme: binding sites for each of the two different sugar

nucleotide precursors (UDP-GlcNAc and UDP-GlcA), two different glycosyltransferase activities, one or more binding sites that anchor the growing HA polymer to the enzyme (perhaps related to a B-X_7-B motif), and a ratchet-like transfer reaction that moves the growing polymer one sugar at a time. This later activity is likely coincident with the stepwise advance of the polymer through the membrane. All of these functions, and perhaps others as yet unknown, are present in a relatively small protein ranging in size from 419 (spHAS) to 588 (xHAS) amino acids.

Although all the available evidence supports the conclusion that only the spHAS protein is required for HA biosynthesis in bacteria or *in vitro*, it is possible that the larger eukaryotic HAS family members are part of multicomponent complexes. Since the eukaryotic HAS proteins are ~40% larger than spHAS, their additional protein domains could be involved in more elaborate functions such as intracellular trafficking and localization, regulation of enzyme activity, and mediating interactions with other cellular components.

The unexpected finding that there are multiple vertebrate HAS genes encoding different synthases strongly supports the emerging consensus that HA is an important regulator of cell behavior and not simply a structural component in tissues. Thus, in less than six months, the field moved from one known, cloned HAS (spHAS) to recognition of a multigene family that promises rapid, numerous, and exciting future advances in our understanding of the synthesis and biology of HA.

For example, disclosed herei ıfter are the sequences of the two HAS genes: from *Pasturella ıultocida;* and (2) *Paramecium bursaria chlorella virus* (PBCV-1 The presence of hyaluronan synthase in these two systems and he purification and use of the hyaluronan synthase from these tw=different systems indicates an ability to purify and isolate r=leic acid sequences encoding enzymatically active hyaluronan= synthase in many different prokaryotic and viral sources.

Group C *Streptococcus equisi=lis* strain D181 synthesizes and secretes hyaluronic acid (HA). In—stigators have used this strain and Group A *Streptococcus pyogene* :rains, such as S43 and A111, to study the biosynthesis of HA ınd to characterize the HA-synthesizing activity in terms of :s divalent cation requirement, precursor (UDP-GlcNAc and UDP-Glc utilization, and optimum pH.

Traditionally, HA has been p=pared commercially by isolation from either rooster combs or extra llular media from Streptococcal cultures. One method which has b=n developed for preparing HA is through the use of cultures of HA-=oducing Streptococcal bacteria. U.S. Patent No. 4,517,295 descri■s such a procedure wherein HA-producing Streptococci are fermen :d under anaerobic conditions in a CO_2-enriched growth medium. Inder these conditions, HA is produced and can be extracted fro=the broth. It is generally felt that isolation of HA from rooster ombs is laborious and difficult, since one starts with HA in a l=s pure state. The advantage of isolation from rooster combs is ■at the HA produced is of higher molecular weight. However, p=paration of HA by bacterial

fermentation is easier, since the HA is of higher purity to start with. Usually, however, the molecular weight of HA produced in this way is smaller than that from rooster combs. Therefore, a technique that would allow the production of high molecular weight HA by bacterial fermentation would be an improvement over existing procedures.

High molecular weight HA has a wide variety of useful applications -- ranging from cosmetics to eye surgery. Due to its potential for high viscosity and its high biocompatibility, HA finds particular application in eye surgery as a replacement for vitreous fluid. HA has also been used to treat racehorses for traumatic arthritis by intra-articular injections of HA, in shaving cream as a lubricant, and in a variety of cosmetic products due to its physiochemical properties of high viscosity and its ability to retain moisture for long periods of time. In fact, in August of 1997 the U.S. Food and Drug Agency approved the use of high molecular weight HA in the treatment of severe arthritis through the injection of such high molecular weight HA directly into the affected joints. In general, the higher molecular weight HA that is employed the better. This is because HA solution viscosity increases with the average molecular weight of the individual HA polymer molecules in the solution. Unfortunately, very high molecular weight HA, such as that ranging up to 10^7, has been difficult to obtain by currently available isolation procedures.

To address these or other difficulties, there is a need for new methods and constructs that can be used to produce HA having

one or more improved properties such as greater purity or ease of preparation. In particular, there is a need to develop methodology for the production of larger amounts of relatively high molecular weight and relatively pure HA than is currently commercially available. There is yet another need to be able to develop methodology for the production of HA having a modified size distribution ($HA_{\Delta size}$) as well as HA having a modified structure ($HA_{\Delta mod}$).

The present invention addresses one or more shortcomings in the art. Using recombinant DNA technology, a purified nucleic acid segment having a coding region encoding enzymatically active seHAS is disclosed and claimed in conjunction, with methods to produce an enzymatically active HA synthase, as well as methods for using the nucleic acid segment in the preparation of recombinant cells which produce HAS and its hyaluronic acid product.

Thus, it is an object of the present invention to provide a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

It is a further object of the present invention to provide a recombinant vector which includes a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

It is still a further object of the present invention to provide a recombinant host cell transformed with a recombinant vector which includes a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

It is yet another object of the present invention to provide a method for detecting a bacterial cell that expresses HAS.

It is another object of the present invention to provide a method for producing high and/or low molecular weight hyaluronic acid from a hyaluronate synthase gene, such as seHAS, as well as methods for producing HA having a modified size distribution and/or a modified structure.

These and other objects of the present invention will become apparent in light of the attached specification, claims, and drawings.

BRIEF SUMMARY OF THE INVENTION

The present invention involves the application of recombinant DNA technology to solving one or more problems in the art of hyaluronic acid (HA) preparation. These problems are addressed through the isolation and use of a nucleic acid segment having a coding region encoding the enzymatically active *Streptococcus equisimilis* (seHAS) hyaluronate synthase gene, a gene responsible for HA chain biosynthesis. The seHAS gene was cloned from DNA of an appropriate microbial source and engineered into useful recombinant constructs for the preparation of HA and for the preparation of large quantities of the HAS enzyme itself.

The present invention encompasses a novel gene, seHAS. The expression of this gene correlates with virulence of *Streptococcal* Group A and Group C strains, by providing a means of escaping phagocytosis and immune surveillance. The terms "hyaluronic acid synthase", "hyaluronate synthase", "hyaluronan synthase" and "HA

synthase", are used interchangeably to describe an enzyme that polymerizes a glycosaminoglycan polysaccharide chain composed of alternating glucuronic acid and N-acetylglucosamine sugars, β 1,3 and β 1,4 linked. The term "seHAS" describes the HAS enzyme derived from *Streptococcus equisimilis*.

The present invention concerns the isolation and characterization of a hyaluronate or hyaluronic acid synthase gene, cDNA, and gene product (HAS), as may be used for the polymerization of glucuronic acid and N-acetylglucosamine into the glycosaminoglycan hyaluronic acid. The present invention identifies the seHAS locus and discloses the nucleic acid sequence which encodes for the enzymatically active seHAS gene from *Streptococcus equisimilis*. The HAS gene also provides a new probe to assess the potential of bacterial specimens to produce hyaluronic acid.

Through the application of techniques and knowledge set forth herein, those of skill in the art will be able to obtain nucleic acid segments encoding the seHAS gene. As those of skill in the art will recognize, in light of the present disclosure, these advantages provide significant utility in being able to control the expression of the seHAS gene and control the nature of the seHAS gene product, the seHAS enzyme, that is produced.

Accordingly, the invention is directed to the isolation of a purified nucleic acid segment which has a coding region encoding enzymatically active HAS, whether it be from prokaryotic or eukaryotic sources. This is possible because the enzyme, and

indeed the gene, is one found in both eukaryotes and some prokaryotes. Eukaryotes are also known to produce HA and thus have HA synthase genes that can be employed in connection with the invention.

HA synthase-encoding nucleic acid segments of the present invention are defined as being isolated free of total chromosomal or genomic DNA such that they may be readily manipulated by recombinant DNA techniques. Accordingly, as used herein, the phrase "a purified nucleic acid segment" refers to a DNA segment isolated free of unrelated chromosomal or genomic DNA and retained in a state rendering it useful for the practice of recombinant techniques, such as DNA in the form of a discrete isolated DNA fragment, or a vector (e.g., plasmid, phage or virus) incorporating such a fragment.

A preferred embodiment of the present invention is a purified nucleic acid segment having a coding region encoding enzymatically active HAS. In particular, the purified nucleic acid segment encodes the seHAS of SEQ ID NO:2 or the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

Another embodiment of the present invention comprises a purified nucleic acid segment having a coding region encoding enzymatically active HAS and the purified nucleic acid segment is capable of hybridizing to the nucleotide sequence of SEQ ID NO:1.

The present invention also comprises a natural or recombinant vector consisting of a plasmid, cosmid, phage, or virus vector.

The recombinant vector may also comprise a purified nucleic acid segment having a coding region encoding enzymatically active HAS.

In particular, the purified nucleic acid segment encodes the seHAS of SEQ ID NO:2 or the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1. If the recombinant vector is a plasmid, it may further comprise an expression vector. The expression vector may also include a promoter operatively linked to the enzymatically active HAS coding region.

In another preferred embodiment, the present invention comprises a recombinant host cel such as a prokaryotic cell transformed with a recombinant vector. The recombinant vector includes a purified nucleic acid segment having a coding region encoding enzymatically active HAS. In particular, the purified nucleic acid segment encodes the seHAS of SEQ ID NO:2 or the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

The present invention also comprises a recombinant host cell, such as an eukaryotic cell transfected with a recombinant vector comprising a purified nucleic acid segment having a coding region encoding enzymatically active HAS In particular, the purified nucleic acid segment encodes the seHAS of SEQ ID NO:2 or the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1. The concept is to create a specifically modified seHAS gene that encodes an enzymatically

active HAS capable of producing a hyaluronic acid polymer having a modified structure or a modified size distribution.

The present invention further comprises a recombinant host cell which is electroporated to introduce a recombinant vector into the recombinant host cell. The recombinant vector may include a purified nucleic acid segment having a coding region encoding enzymatically active HAS. In particular, the purified nucleic acid segment encodes the seHAS of SEQ ID NO:2 or the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1. The enzymatically active HAS may also be capable of producing a hyaluronic acid polymer having a modified structure or a modified size distribution.

In yet another preferred embodiment, the present invention comprises a recombinant host cell which is transduced with a recombinant vector which includes a purified nucleic acid segment having a coding region encoding enzymatically active HAS. In particular, the purified nucleic acid segment encodes the seHAS of SEQ ID NO:2 or the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1. The enzymatically active HAS is also capable of producing a hyaluronic acid polymer having a modified structure or a modified size distribution.

The present invention also comprises a purified composition, wherein the purified composition comprises a polypeptide having a coding region encoding enzymatically active HAS and further having an amino acid sequence in accordance with SEQ ID NO:2.

In another embodiment, the invention comprises a method for detecting a DNA species, comprising the steps of: (1) obtaining a DNA sample; (2) contacting the DNA sample with a purified nucleic acid segment in accordance with SEQ ID NO:1; (3) hybridizing the DNA sample and the purified nucleic acid segment thereby forming a hybridized complex; and (4) detecting the complex.

The present invention also comprises a method for detecting a bacterial cell that expresses mRNA encoding seHAS, comprising the steps of: (1) obtaining a bacterial cell sample; (2) contacting at least one nucleic acid from the bacterial cell sample with purified nucleic acid segment in accordance with SEQ ID NO:1; (3) hybridizing the at least one nucleic acid and the purified nucleic acid segment thereby forming a hybridized complex; and (4) detecting the hybridized complex wherein the presence of the hybridized complex is indicative of a bacterial strain that expresses mRNA encoding seHAS.

The present invention also comprises methods for detecting the presence of either seHAS or spHAS in a cell. In particular, the method comprises using the oligonucleotides set forth in Seq. ID Nos.: 3-8 as probes. These oligonucleotides would a allow a practitioner to search and detect the presence of seHAS or spHAS in a cell.

The present invention further comprises a method for producing hyaluronic acid, comprising the steps of: (1) introducing a purified nucleic acid segment having a coding region encoding enzymatically active HAS into a host organism, wherein the host

organism contains nucleic acid segments encoding enzymes which produce UDP-GlcNAc and UDP-GlcA; (2) growing the host organism in a medium to secrete hyaluronic acid; and (3) recovering the secreted hyaluronic acid.

The method may also include the step of extracting the secreted hyaluronic acid from the medium as well as the step of purifying the extracted hyaluronic acid. Furthermore, the host organism may secrete a structurally modified hyaluronic acid or a size modified hyaluronic acid.

The present invention further comprises a pharmaceutical composition comprising a preselected pharmaceutical drug and an effective amount of hyaluronic acid produced by a recombinant HAS. The pharmaceutical composition may have a hyaluronic acid having a modified molecular weight pharmaceutical composition capable of evading an immune response. The modified molecular weight may also produce a pharmaceutical composition capable of targeting a specific tissue or cell type within the patient having an affinity for the modified molecular weight pharmaceutical composition.

The present invention also comprises a purified and isolated nucleic acid sequence encoding enzymatically active seHAS, where the nucleic acid sequence is (a) the nucleic acid sequence in accordance with SEQ ID NO:1; (b) complementary nucleic acid sequences to the nucleic acid sequence in accordance with SEQ ID NO:1; (c) nucleic acid sequences which will hybridize to the nucleic acid in accordance with SEQ ID NO:1; and (d) nucleic acid

sequences which will hybridize to ıe complementary nucleic acid sequences of SEQ ID NO:1.

The present invention furth■r comprises a purified and isolated nucleic acid segment cons=ting essentially of a nucleic acid segment encoding enzymaticall=active HAS.

The present invention also com-ises an isolated nucleic acid segment consisting essentially of =nucleic acid segment encoding seHAS having a nucleic acid segment ufficiently duplicative of the nucleic acid segment in accordar=e of SEQ ID NO:1 to allow possession of the biological p=perty of encoding for an enzymatically active HAS. The nuc ic acid segment may also be a cDNA sequence.

The present invention also co rises a purified nucleic acid segment having a coding region enc ling enzymatically active HAS, wherein the purified nucleic acid s ment is capable of hybridizing to the nucleotide sequence in acco=dance with SEQ ID NO:1.

BRIEF DESCRIPTION OF THE SEV■AL VIEWS OF THE DRAWINGS

FIG. 1 depicts that cross h=ridization between seHAS and spHAS genes does not occur.

FIG. 2 figuratively depicts —e relatedness of seHAS to the bacterial and eukaryotic HAS prote=s.

FIG. 3 figuratively depicts =olutionary relationships among some of the known hyaluronan synth=ses.

FIG. 4 depicts the HA size d=stribution produced by various engineered Streptococcal HAS enzy=s.

FIG. 5 figuratively depicts the overexpression of recombinant seHAS and spHAS in E. coli.

FIG. 6 depicts purification of Streptococcal HA synthase.

FIG. 7 depicts a gel filtration analysis of HA synthesized by recombinant streptococcal HAS expressed in yeast membranes.

FIG. 8 is a Western blot analysis of recombinant seHAS using specific antibodies.

FIG. 9 is a kinetic analysis of the HA size distributions produced by recombinant seHAS and spHAS.

FIG. 10 graphically depicts the hydropathy plots for seHAS and predicted membrane associated regions.

FIG. 11 is a graphical model for the topologic organization of seHAS in the membrane.

FIG. 12 is a demonstration of the synthesis of authentic HA by the recombinant seHAS.

FIG. 13 depicts the recognition of nucleic acid sequences encoding seHAS, encoding spHAS, or encoding both seHAS and spHAS using specific oligonucleotides and PCR.

FIG. 14 depicts oligonucleotides used for specific PCR hybridization.

DETAILED DESCRIPTION OF THE INVENTION

Before explaining at least one embodiment of the invention in detail, it is to be understood that the invention is not limited in its application to the details of construction and the arrangement of the components set forth in the following description or illustrated in the drawings. The invention is capable of other

embodiments or of being practiced or carried out in various ways. Also, it is to be understood that the phraseology and terminology employed herein is for purpose of description and should not be regarded as limiting.

As used herein, the term "nucleic acid segment" and "DNA segment" are used interchangeably and refer to a DNA molecule which has been isolated free of total genomic DNA of a particular species. Therefore, a "purified" DNA or nucleic acid segment as used herein, refers to a DNA segment which contains a Hyaluronate Synthase ("HAS") coding sequence yet is isolated away from, or purified free from, unrelated genomic DNA, for example, total *Streptococcus equisimilis* or, for example, mammalian host genomic DNA. Included within the term "DNA segment", are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phage, viruses, and the like.

Similarly, a DNA segment comprising an isolated or purified seHAS gene refers to a DNA segment including HAS coding sequences isolated substantially away from other naturally occurring genes or protein encoding sequences. In this respect, the term "gene" is used for simplicity to refer to a functional protein, polypeptide or peptide encoding unit. As will be understood by those in the art, this functional term includes genomic sequences, cDNA sequences or combinations thereof. "Isolated substantially away from other coding sequences" means that the gene of interest, in this case seHAS, forms the significant part of the coding region of

the DNA segment, and that the DNA segment does not contain large portions of naturally-occurring coding DNA, such as large chromosomal fragments or other functional genes or DNA coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to, or intentionally left in the segment by the hand of man.

Due to certain advantages associated with the use of prokaryotic sources, one will likely realize the most advantages upon isolation of the HAS gene from prokaryotes such as *S. pyogenes*, *S. equisimilis*, or *P. multocida*. One such advantage is that, typically, eukaryotic enzymes may require significant post-translational modifications that can only be achieved in a eukaryotic host. This will tend to limit the applicability of any eukaryotic HA synthase gene that is obtained. Moreover, those of ordinary skill in the art will likely realize additional advantages in terms of time and ease of genetic manipulation where a prokaryotic enzyme gene is sought to be employed. These additional advantages include (a) the ease of isolation of a prokaryotic gene because of the relatively small size of the genome and, therefore, the reduced amount of screening of the corresponding genomic library and (b) the ease of manipulation because the overall size of the coding region of a prokaryotic gene is significantly smaller due to the absence of introns. Furthermore, if the product of the seHAS gene (i.e., the enzyme) requires posttranslational modifications, these would best be achieved in a similar

prokaryotic cellular environment ost) from which the gene was derived.

Preferably, DNA sequences i accordance with the present invention will further include gen ic control regions which allow the expression of the sequence in elected recombinant host. Of course, the nature of the control egion employed will generally vary depending on the particul use (e.g., cloning host) envisioned.

In particular embodiments, th nvention concerns isolated DNA segments and recombinant vectors i orporating DNA sequences which encode a seHAS gene, that include within its amino acid sequence an amino acid sequence in accorda with SEQ ID NO:2. Moreover, in other particular embodiments, e invention concerns isolated DNA segments and recombinant vect s incorporating DNA sequences which encode a gene that includes ithin its amino acid sequence the amino acid sequence of an HAS ne or DNA, and in particular to an HAS gene or cDNA, correspondi to *Streptococcus equisimilis* HAS. For example, where the DNA gment or vector encodes a full length HAS protein, or is intende for use in expressing the HAS protein, preferred sequences are t se which are essentially as set forth in SEQ ID NO:2.

Nucleic acid segments havi HA synthase activity may be isolated by the methods describe herein. The term "a sequence essentially as set forth in SEQ NO:2" means that the sequence substantially corresponds to a rtion of SEQ ID NO:2 and has relatively few amino acids whi are not identical to, or a

biologically functional equivalent of, the amino acids of SEQ ID NO:2. The term "biologically functional equivalent" is well understood in the art and is further defined in detail herein, as a gene having a sequence essentially as set forth in SEQ ID NO:2, and that is associated with the ability of prokaryotes to produce HA or a hyaluronic acid coat.

For instance, the seHAS and spHAS coding sequences are approximately 70% identical and rich in the bases adenine (A) and thymine (T). SeHAS base content is A-26.71%, C-19.13%, G-20.81%, and T-33.33% (A/T = 60%). Whereas spHAS is A-31.34%, C-16.42%, G-16.34%, and T-35.8% (A/T = 67%). Those of ordinary skill in the art would be surprised that the seHAS coding sequence does not hybridize with the spHAS gene and vice versa, despite their being 70% identical. This unexpected inability to cross-hybridize could be due to short interruptions of mismatched bases throughout the open reading frames. The inability of spHAS and seHAS to cross-hybridize is shown in FIG. 1. The longest stretch of identical nucleotides common to both the seHAS and the spHAS coding sequences is only 20 nucleotides. In addition, the very A-T rich sequences will form less stable hybridization complexes than G-C rich sequences. Another possible explanation could be that there are several stretches of As or Ts in both sequences that could hybridize in a misaligned and unstable manner. This would put the seHAS and spHAS gene sequences out of frame with respect to each other, thereby decreasing the probability of productive hybridization.

Because of this unique phenomena of two genes encoding proteins which are 70% identical not being capable of cross-hybridizing to one another, it is beneficial to think of the claimed nucleic acid segment in terms of its function; i.e. a nucleic acid segment which encodes enzymatically active hyaluronate synthase. One of ordinary skill in the art would appreciate that a nucleic acid segment encoding enzymatically active hyaluronate synthase may contain conserved or semi-conserved substitutions to the sequences set forth in SEQ ID NOS: 1 and 2 and yet still be within the scope of the invention.

In particular, the art is replete with examples of practitioners ability to make structural changes to a nucleic acid segment (i.e. encoding conserved or semi-conserved amino acid substitutions) and still preserve its enzymatic or functional activity. See for example:) Risler et al. "Amino Acid Substitutions in Structurally Related Proteins. A Pattern Recognition Approach." J. Mol. ol. 204:1019-1029 (1988) ["... according to the observed exchangeability of amino acid side chains, only four groups could be elineated; (i) Ile and Val; (ii) Leu and Met, (iii) Lys, Arg, and ln, and (iv) Tyr and Phe."]; (2) Niefind et al. "Amino Acid Similarity Coefficients for Protein Modeling and Sequence Alignment Derived from Main-Chain Folding Anoles." J. Mol. Biol. 219:481-7 (1991) [similarity parameters allow amino acid substitutions to be designed]; and (3) Overington et al. "Environment-Specific Amino Acid Substitution Tables: Tertiary Templates and Prediction of Protein Folds," Protein

Science 1:216-226 (1992) ["Analysis of the pattern of observed substitutions as a function of local environment shows that there are distinct patterns..." Compatible changes can be made.]

These references and countless others, indicate that one of ordinary skill in the art, given a nucleic acid sequence, could make substitutions and changes to the nucleic acid sequence without changing its functionality. Also, a substituted nucleic acid segment may be highly identical and retain its enzymatic activity with regard to its unadulterated parent, and yet still fail to hybridize thereto.

The invention discloses nucleic acid segments encoding enzymatically active hyaluronate synthase - seHAS and spHAS. Although seHAS and spHAS are 70% identical and both encode enzymatically active hyaluronate synthase, they do not cross hybridize. Thus, one of ordinary skill in the art would appreciate that substitutions can be made to the seHAS nucleic acid segment listed in SEQ ID NO: 1 without deviating outside the scope and claims of the present invention. Standardized and accepted functionally equivalent amino acid substitutions are presented in Table I.

TABLE

Amino Acid Group	Conservative and Semi-conservative Substitutions
NonPolar R Groups	Alanine, Valine, Leucine, Isoleucine, Proline, Methionine, Phenylalanine, Tryptophan
Polar, but uncharged, R Groups	Glycine, Serine, Threonine, Cysteine, Asparagine, Glutamine
Negatively Charged R Groups	Aspartic Acid, Glutamic Acid
Positively Charged R Groups	Lysine, Arginine, Histidine

Another preferred embodiment of the present invention is a purified nucleic acid segment that encodes a protein in accordance with SEQ ID NO:2, further defined as a recombinant vector. As used herein, the term "recombinant vector" refers to a vector that has been modified to contain a nucleic acid segment that encodes an HAS protein, or fragment thereof. The recombinant vector may be further defined as an expression vector comprising a promoter operatively linked to said HAS encoding nucleic acid segment.

A further preferred embodiment of the present invention is a host cell, made recombinant with a recombinant vector comprising an HAS gene. The preferred recombinant host cell may be a prokaryotic cell. In another embodiment, the recombinant host cell is a eukaryotic cell. As used herein, the term "engineered" or "recombinant" cell is intended to refer to a cell into which a recombinant gene, such as a gene encoding HAS, has been introduced. Therefore, engineered cells are distinguishable from naturally occurring cells which do not contain a recombinantly introduced gene. Engineered cells are thus cells having a gene or genes

introduced through the hand of man. Recombinantly introduced genes will either be in the form of a cDNA gene, a copy of a genomic gene, or will include genes positioned adjacent to a promoter not naturally associated with the particular introduced gene.

Where one desires to use a host other than *Streptococcus*, as may be used to produce recombinant HA synthase, it may be advantageous to employ a prokaryotic system such as *E. coli*, *B. subtilis*, *Lactococcus sp.*, or even eukaryotic systems such as yeast or Chinese hamster ovary, African green monkey kidney cells, VERO cells, or the like. Of course, where this is undertaken it will generally be desirable to bring the HA synthase gene under the control of sequences which are functional in the selected alternative host. The appropriate DNA control sequences, as well as their construction and use, are generally well known in the art as discussed in more detail hereinbelow.

In preferred embodiments, the HA synthase-encoding DNA segments further include DNA sequences, known in the art functionally as origins of replication or "replicons", which allow replication of contiguous sequences by the particular host. Such origins allow the preparation of extrachromosomally localized and replicating chimeric segments or plasmids, to which HA synthase DNA sequences are ligated. In more preferred instances, the employed origin is one capable of replication in bacterial hosts suitable for biotechnology applications. However, for more versatility of cloned DNA segments, it may be desirable to alternatively or even

additionally employ origins recogni_ed by other host systems whose use is contemplated (such as in a =uttle vector).

The isolation and use of other eplication origins such as the SV40, polyoma or bovine papillom=virus origins, which may be employed for cloning or expression = a number of higher organisms, are well known to those of ordinar=skill in the art. In certain embodiments, the invention may t■s be defined in terms of a recombinant transformation vector ıich includes the HA synthase coding gene sequence together wi_ı an appropriate replication origin and under the control of se :cted control regions.

Thus, it will be appreciated b=those of skill in the art that other means may be used to obtain t : HAS gene or cDNA, in light of the present disclosure. For examp■, polymerase chain reaction or RT-PCR produced DNA fragments may e obtained which contain full complements of genes or cDNAs from■ number of sources, including other strains of *Streptococcus* or ■om eukaryotic sources, such as cDNA libraries. Virtually any mo=cular cloning approach may be employed for the generation of DNA :agments in accordance with the present invention. Thus, the on ' limitation generally on the particular method employed for DN=isolation is that the isolated nucleic acids should encode a biolc=ically functional equivalent HA synthase.

Once the DNA has been isolat= it is ligated together with a selected vector. Virtually any c=ning vector can be employed to realize advantages in accordance=with the invention. Typical useful vectors include plasmids a= phages for use in prokaryotic

organisms and even viral vectors for use in eukaryotic organisms. Examples include pKK223-3, pSA3, recombinant lambda, SV40, polyoma, adenovirus, bovine papilloma virus and retroviruses. However, it is believed that particular advantages will ultimately be realized where vectors capable of replication in both *Lactococcus* or *Bacillus* strains and *E. coli* are employed.

Vectors such as these, exemplified by the pSA3 vector of Dao and Ferretti or the pAT19 vector of Trieu-Cuot, et al., allow one to perform clonal colony selection in an easily manipulated host such as *E. coli*, followed by subsequent transfer back into a food grade *Lactococcus* or *Bacillus* strain for production of HA. These are benign and well studied organisms used in the production of certain foods and biotechnology products. These are advantageous in that one can augment the *Lactococcus* or *Bacillus* strain's ability to synthesize HA through gene dosaging (i.e., providing extra copies of the HA synthase gene by amplification) and/or inclusion of additional genes to increase the availability of HA precursors. The inherent ability of a bacterium to synthesize HA can also be augmented through the formation of extra copies, or amplification, of the plasmid that carries the HA synthase gene. This amplification can account for up to a 10-fold increase in plasmid copy number and, therefore, the HA synthase gene copy number.

Another procedure that would further augment HA synthase gene copy number is the insertion of multiple copies of the gene into the plasmid. Another technique would include integrating the HAS

gene into chromosomal DNA. This extra amplification would be especially feasible, since the bacterial HA synthase gene size is small. In some scenarios, the chromosomal DNA-ligated vector is employed to transfect the host that is selected for clonal screening purposes such as *E. coli* through the use of a vector that is capable of expressing the inserted DNA in the chosen host.

Where a eukaryotic source such as dermal or synovial fibroblasts or rooster comb cells is employed, one will desire to proceed initially by preparing a cDNA library. This is carried out first by isolation of mRNA from the above cells, followed by preparation of double stranded cDNA using an enzyme with reverse transcriptase activity and ligation with the selected vector. Numerous possibilities are available and known in the art for the preparation of the double stranded cDNA, and all such techniques are believed to be applicable. A preferred technique involves reverse transcription. Once a population of double stranded cDNAs is obtained, a cDNA library is prepared in the selected host by accepted techniques, such as by ligation into the appropriate vector and amplification in the appropriate host. Due to the high number of clones that are obtained, and the relative ease of screening large numbers of clones by the techniques set forth herein, one may desire to employ phage expression vectors, such as λgt11, λgt12, λGem11, and/or λZAP for the cloning and expression screening of cDNA clones.

In certain other embodiments the invention concerns isolated DNA segments and recombinant vectors that include within their

sequence a nucleic acid sequence essentially as set forth in SEQ ID NO:1. The term "essentially as set forth in SEQ ID NO:1" is used in the same sense as described above and means that the nucleic acid sequence substantially corresponds to a portion of SEQ ID NO:1, and has relatively few codons which are not identical, or functionally equivalent, to the codons of SEQ ID NO:1. The term "functionally equivalent codon" is used herein to refer to codons that encode the same amino acid, such as the six codons for arginine or serine, as set forth in Table I, and also refers to codons that encode biologically equivalent amino acids.

It will also be understood that amino acid and nucleic acid sequences may include additional residues, such as additional N- or C-terminal amino acids or 5' or 3' nucleic acid sequences, and yet still be essentially as set forth in one of the sequences disclosed herein, so long as the sequence meets the criteria set forth above, including the maintenance of biological protein activity where protein expression and enzyme activity is concerned. The addition of terminal sequences particularly applies to nucleic acid sequences which may, for example, include various non-coding sequences flanking either of the 5' or 3' portions of the coding region or may include various internal sequences, which are known to occur within genes. In particular, the amino acid sequence of the HAS gene in eukaryotes appears to be 40% larger than that found in prokaryotes.

Allowing for the degeneracy of the genetic code as well as conserved and semi-conserved substitutions, sequences which have

between about 40% and about 80%; or more preferably, between about 80% and about 90%; or even more preferably, between about 90% and about 99%; of nucleotides which are identical to the nucleotides of SEQ ID NO:1 will be sequences which are "essentially as set forth in SEQ ID NO:1". Sequences which are essentially the same as those set forth in SEQ ID NO:1 may also be functionally defined as sequences which are capable of hybridizing to a nucleic acid segment containing the complement of SEQ ID NO:1 under standard or less stringent hybridizing conditions. Suitable standard hybridization conditions will be well known to those of skill in the art and are clearly set forth herein.

The term "standard hybridization conditions" as used herein, is used to describe those conditions under which substantially complementary nucleic acid segments will form standard Watson-Crick base-pairing. A number of factors are known that determine the specificity of binding or hybridization, such as pH, temperature, salt concentration, the presence of agents, such as formamide and dimethyl sulfoxide, the length of the segments that are hybridizing, and the like. When it is contemplated that shorter nucleic acid segments will be used for hybridization, for example fragments between about 14 and about 100 nucleotides, salt and temperature preferred conditions for hybridization will include 1.2-1.8 x HPB at 40-50°C.

Naturally, the present invention also encompasses DNA segments which are complementary, or essentially complementary, to the sequence set forth in SEQ ID NO:1. Nucleic acid sequences which

are "complementary" are those which are capable of base-pairing according to the standard Watson-Crick complementarity rules. As used herein, the term "complementary sequences" means nucleic acid sequences which are substantially complementary, as may be assessed by the same nucleotide comparison set forth above, or as defined as being capable of hybridizing to the nucleic acid segment of SEQ ID NO:1.

The nucleic acid segments of the present invention, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, epitope tags, poly histidine regions, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol.

Naturally, it will also be understood that this invention is not limited to the particular nucleic acid and amino acid sequences of SEQ ID NO:1 and 2. Recombinant vectors and isolated DNA segments may therefore variously include the HAS coding regions themselves, coding regions bearing selected alterations or modifications in the basic coding region, or they may encode larger polypeptides which nevertheless include HAS-coding regions or may encode biologically functional equivalent proteins or peptides which have variant amino acids sequences.

For instance, we have found, characterized, and purified hyaluronate synthase in two other systems: (a) the gram-negative bacteria *Pasturella multocida* (SEQ ID NO:19); and (2) chlorella virus PBCV-1 (SEQ ID NOS:7 and 8). The presence of hyaluronan synthase in these two systems and our ability to purify and use the hyaluronan synthase from these two different systems indicates our ability to purify and isolate nucleic acid sequences encoding enzymatically active hyaluronan synthase.

The capsule of Carter Type A *P. multocida* (SEQ ID NO:19) was long suspected of containing hyaluronic acid-HA. Characterization of the HA synthase of *P. multocida* led to interesting enzymological differences between it and the seHAS and spHAS proteins.

P. multocida cells produce a readily visible extracellular HA capsule, and since the two streptococcal HASs are membrane proteins, membrane preparations of the fowl cholera pathogen were tested. In early trials, crude membrane fractions derived from ultrasonication alone possessed very low levels of UDP-GlcNAc-dependent UDP-[^{14}C]GlcA incorporation into HA[~0.2 pmol of GlcA transfer (μg of proteins)$^{-1}$h^{-1}] when assayed under conditions similar to those for measuring streptococcal HAS activity. The enzyme from *E. coli* with the recombinant *hasA* plasmid was also recalcitrant to isolation at first. These results were in contrast to the easily detectable amounts obtained from *Streptococcus* by similar methods.

An alternative preparation protocol using ice-cold lysozyme treatment in the presence of protease inhibitors in conjunction with ultrasonication allowed the substantial recovery of HAS

activity from both species of Gram-negative bacteria. Specific activities for HAS of 5-10 pmol of GlcA transferred (μg of protein)$^{-1}h^{-1}$ were routinely obtained for crude membranes of wild-type *P. multocida* with the new method. In the absence of UDP-GlcNAc, virtually no radioactivity (<1% of identical assay with both sugar precursors) from UDP-[^{14}C]GlcA was incorporated into higher molecular weight material. Membranes prepared from the acapsular mutant, TnA, possessed no detectable HAS activity when supplemented with both sugar nucleotide precursors (data not shown). Gel-filtration analysis using a Sephacryl S-200 column indicates that the molecular mass of the majority of the ^{14}C-labeled product synthesized *in vitro* is ≥8 x 10^4 Da since the material elutes in the void volumes, such a value corresponds to a HA molecule composed of at least 400 monomers. This product is sensitive to *Streptomyces* hyaluronidase digestion but resistant to protease treatment.

The parameters of the HAS assay were varied to maximize incorporation of UDP-sugars into polysaccharide by *P. multocida* membranes. Streptococcal spHAS requires Mg^{2+} and therefore this metal ion was included in the initial assays of *P. multocida* membranes. The *P. multocida* HAS (pmHAS) was relatively active from pH 6.5 to 8.6 in Tris-type buffers with an optimum at pH 7. The HAS activity was linear with respect to the incubation time at neutral pH for at least 1 h. The pmHAS was apparently less active at higher ionic strengths because the addition of 100 mM NaCl to

the reaction containing 50 mM Tris pH 7, and 20 mM $MgCl_2$ reduced sugar incorporation by ~50%.

The metal ion specificity of the pmHAS was assessed at pH 7. Under metal-free conditions in the presence of EDTA, no incorporation of radiolabeled precursor into polysaccharide was detectable (<0.5% of maximal signal). Mn^{2+} gave the highest incorporation rates at the lowest ion concentrations for the tested metals (Mg, Mn, Co, Cu, and Ni). Mg^{2+} gave about 50% of the Mn^{2+} stimulation but at 10-fold higher concentrations. Co^{2+} or Ni^{2+} at 10mM supported lower levels of activity (20% or 9%, respectively, of 1 mM Mn^{2+} assays), but membranes supplied with 10 mM Cu^{2+} were inactive. Indeed, mixing 10 mM Cu^{2+} and 20 mM Mg^{2+} with the membrane preparation resulted in almost no incorporation of label into polysaccharide (<0.8% of Mg only value).

Initial characterization of the pmHAS was performed in the presence of Mg^{2+}. The binding affinity of the enzyme for its sugar nucleotide precursors was assessed by measuring the apparent K_M value. Incorporation of [^{14}C]GlcA or [^{3}H]GlcNAc into polysaccharide was monitored at varied concentrations of UDP-GlcNAc or UDP-GlcA, respectively. In Mg^{2+}-containing buffers, the apparent K_M values of ~20 μM for UDP-GlcA and ~75 μM for UDP-GlcNAc were determined utilizing Hanes-Woolf plots ([S] / *versus* [S]) of the titration data. The V_{max} values for both sugars were the same because the slopes, corresponding to $1/V_{max}$, of the Hanes-Woolf plots were equivalent. In comparison to results from assays with Mg^{2+}, the K_M

value for UDP-GlcNAc was increased by about 25-50% to ~105 μM and the V_{max} increased by a factor of 2-3-fold in the presence of Mn^{2+}.

The HA synthase enzymes from either *P. multocida*, *S. equisimilis*, or *S. pyogenes* utilizes UDP-sugars, but they possess somewhat different kinetic optima with respect to pH and metal ion dependence and K_M values. The enzymes are most active at pH 7; however, the pmHAS reportedly displays more activity at slightly acidic pH and is relatively inactive above pH 7.4. The pmHAS utilizes Mn^{2+} more efficiently than Mg^{2+} under the *in vitro* assay conditions, but the identity of the physiological metal cofactor in the bacterial cell is unknown. In comparison, in previous studies with the streptococcal enzyme, Mg^{2+} was much better than Mn^{2+} but the albeit smaller effect of Mn^{2+} was maximal at ~10-fold lower concentrations than the optimal Mg^{2+} concentration. The pmHAS apparently binds the UDP-sugars more tightly than spHAS. The measured K_M values for the pmHAS in crude membranes are about 2-3-fold lower for each substrate than those obtained from the HAS found in streptococcal membranes: 50 or 39 μM for UDP-GlcA and 500 or 150 μM for UDP-GlcNAc, respectively.

By kinetic analyses, the V_{max} of the pmHAS was 2-3-fold higher in the presence of Mn^{2+} than Mg^{2+}, but the UDP-GlcNAc K_M value was increased slightly in assays with the former ion. This observation of apparent lowered affinity suggests that the increased polymerization rate was *not* due to better binding of the Mn^{2+} ion/sugar nucleotide complex to the enzyme active site(s). Therefore, it is possible that Mn^{2+} enhances some other reaction

step, alters another site/structure of the enzyme, or modifies the phospholipid membrane environment. The gene sequence and the protein sequence of pmHAS are shown in SEQ ID NO:19.

Chlorella virus PBCV-1 encodes a functional glycosyltransferase that can synthesize a polysaccharide, hyaluronan [hyaluronic acid, HA]. This finding is contrary to the general observation that viruses either: (a) utilize host cell glycosyltransferases to create new carbohydrate structures, or (b) accumulate host cell glycoconjugates during virion maturation. Furthermore, HA has been generally regarded as restricted to animals and a few of their virulent bacterial pathogens. Though many plant carbohydrates have been characterized, neither HA nor a related analog has previously been detected in cells of plants or protists.

The vertebrate HAS enzymes (DG42, HAS1, HAS2, HAS3) and streptococcal HasA enzymes (spHAS and seHAS) have several regions of sequence similarity. While sequencing the double-stranded DNA genome of virus PBCV-1 [*Paramecium bursaria* chlorella virus], an ORF [open reading frame], A98R (Accession #442580), encoding a 567 residue protein with 28 to 33% amino acid identity to the various HASs was discovered. This protein is designated cvHAS (chlorella virus HA synthase). The gene sequence encoding PBCV-1 and its protein sequence are shown in SEQ ID NOS:7 and 8.

PBCV-1 is the prototype of a family (Phycodnarviridae) of large (175-190 nm diameter) polyhedral, plaque-forming viruses that replicate in certain unicellular, eukaryotic chlorella-like green

algae. PBCV-1 virions contain at least 50 different proteins and a lipid component located inside the outer glycoprotein capsid. The PBCV-1 genome is a linear, nonpermuted 330-kb dsDNA molecule with covalently closed hairpin ends.

Based on its deduced amino acid sequence, the *A98R* gene product should be an integral membrane protein. To test this hypothesis, recombinant A98R was produced in *Escherichia coli* and the membrane fraction was assayed for HAS activity. UDP-GlcA and UDP-GlcNAc were incorporated into the polysaccharide by the membrane fraction derived from cells containing the A98R gene on a plasmid, pCVHAS, (average specific activity 2.5 pmoles GlcA transfer/μg protein/min) but <u>not</u> by samples from control cells (<0.001 pmoles GlcA transfer/μg protein/min). No activity was detected in the soluble fraction of cells transformed with pCVHAS. UDP-GlcA and UDP-GlcNAc were simultaneously required for polymerization. The activity was optimal in Hepes buffer at pH 7.2 in the presence of 10 mM $MnCl_2$, whereas no activity was detected if the metal ion was omitted. Mg^{2+} and Co^{2+} were ~20% as effective as Mn^{2+} at similar concentrations. The pmHAS has a similar metal requirement, but other HASs prefer Mg^{2+}.

The recombinant A98R enzyme synthesized a polysaccharide with an average molecular weight of $3\text{-}6\times10^6$ Da which is smaller than that of the HA synthesized by recombinant spHAS or DG42 xlHAS *in vitro* (~10^7 Da and ~$5\text{-}8\times10^6$ Da, respectively; *13,15*). The polysaccharide was completely degraded by *Streptomyces hyaluroniticus* HA lyase, an enzyme that depolymerizes HA, but not

structurally related glycosaminoglycans such as heparin and chondroitin.

PBCV-1 infected chlorella cells were examined for *A98R* gene expression. A ~1,700-nucleotide *A98R* transcript appeared at ~15 min post-infection and disappeared by 60 min after infection indicating that *A98R* is an early gene. Consequently, membrane fractions from uninfected and PBCV-1 infected chlorella cells were assayed at 50 and 90 min post-infection for HAS activity. Infected cells, but not uninfected cells, had activity. Like the bacterially derived recombinant A98R enzyme, radiolabel incorporation from UDP-[^{14}C]GlcA into polysaccharide depended on both Mn^{2+} and UDP-GlcNAc. This radiolabeled produce was also degraded by HA lyase. Disrupted PBCV-1 virions had no HAS activity.

PBCV-1 infected chlorella cells were analyzed for HA polysaccharide using a highly specific ^{125}I-labeled HA-binding protein. Extracts from cells at 50 and 90 min post-infection contained substantial amounts of HA, but not extracts from uninfected algae or disrupted PBCV-1 virions. The labeled HA-binding protein also interacted with intact infected cells at 50 and 90 min post-infection, but not healthy cells. Therefore, a considerable portion of the newly synthesized HA polysaccharide was immobilized at the outer cell surface of the infected algae. The extracellular HA does not play an obvious role in the interaction between the virus and its algal host because neither plaque size nor plaque number was altered by including either testicular

hyaluronidase (465 units/ml) or free HA polysaccharide (100 μg/ml) in the top agar of the PBCV-1 plaque assay.

The PBCV-1 genome also has additional genes that encode for an UDP-Glc dehydrogenase (UDP-Glc DH) and a glutamine:fructose-6-phosphate aminotransferase (GFAT). UDP-Glc DH converts UDP-Glc into UDP-GlcA, a required precursor for HA biosynthesis. GFAT converts fructose-6-phosphate into glucosamine-6-phosphate, an intermediate in the UDP-GlcNAc metabolic pathway. Both of these PBCV-1 genes, like the *A98R* HAS, are expressed early in infection and encode enzymatically active proteins. The presence of multiple enzymes in the HA biosynthesis pathway indicates that HA production must serve an important function in the life cycle of the chlorella viruses.

HA synthases of *Streptococcus*, vertebrates, and PBCV-1 possess many motifs of 2 to 4 residues that occur in the same relative order. These conserved motifs probably reflect domains crucial for HA biosynthesis as shown in FIG. 2. The protein sequences of Group C seHAS, Group A spHAS, murine HAS1, HAS2, HAS3, and frog HAS are shown aligned in FIG. 2. The alignment of FIG. 2 was accomplished using the DNAsis multiple alignment program. Residues in seHAS identical in other known HAS family members (including human HAS1 and 2, not shown) are denoted by shading and asterisks. The amino acids indicated by dots are conserved in all members of the larger β-glycosyl transferase family. The diamond symbol indicates the highly conserved cysteine residue that may be critical for enzyme activity. The approximate mid-points of predicted membrane domains

MD1 through MD7 are indicated with arrows. X1 indicates *Xenopus laevis*, and MM denotes *Mus musculus*.

Regions of similarity between HASs and other enzymes that synthesize β-linked polysaccharides from UDP-sugar precursors are also being discovered as more glycosyltransferases are sequenced. Examples include bacterial cellulose synthase, fungal and bacterial chitin synthases, and the various HASs. The significance of these similar structural motifs will become more apparent as the three-dimensional structures of glycosyltransferases accumulate.

FIG. 3 depicts the evolutionary relationships among the known hyaluronan synthase. The phylogenetic tree of FIG. 3 was generated by the Higgins-Sharp algorithm using the DNAsis multiple alignment program. The calculated matching percentages are indicated at each branch of the dendrogram.

The DNA segments of the present invention encompass biologically functional equivalent HAS proteins and peptides. Such sequences may arise as a consequence of codon redundancy and functional equivalency which are known to occur naturally within nucleic acid sequences and the proteins thus encoded. Alternatively, functionally equivalent proteins or peptides may be created via the application of recombinant DNA technology, in which changes in the protein structure may be engineered, based on considerations of the properties of the amino acids being exchanged. Changes designed by man may be introduced through the application of site-directed mutagenesis techniques, e.g., to introduce improvements to the enzyme activity or to antigenicity of the HAS

protein or to test HAS mutants in order to examine HA synthase activity at the molecular level.

Also, specific changes to the HAS coding sequence can result in the production of HA having a modified size distribution or structural configuration. One of ordinary skill in the art would appreciate that the HAS coding sequence can be manipulated in a manner to produce an altered hyaluronate synthase which in turn is capable of producing hyaluronic acid having differing polymer sizes and/or functional capabilities. For example, the HAS coding sequence may be altered in such a manner that the hyaluronate synthase has an altered sugar substrate specificity so that the hyaluronate synthase creates a new hyaluronic acid-like polymer incorporating a different structure such as a previously unincorporated sugar or sugar derivative. This newly incorporated sugar could result in a modified hyaluronic acid having different functional properties, a hyaluronic acid having a smaller or larger polymer size/molecular weight, or both. As will be appreciated by one of ordinary skill in the art given the HAS coding sequences, changes and/or substitutions can be made to the HAS coding sequence such that these desired property and/or size modifications can be accomplished. Table II lists sugar nucleotide specificity and magnesium ion requirement of recombinant seHAS.

TABLE I
Sugar nucleotide specificity and
Magnesium ion requirement of recombinant seHAS

Second Sugar nucleotide present (μM)	HA Synthesis* UDP-[^{14}C]GlcA dpm (%)	UDP-[^{3}H]GlcNAc dpm (%)
None	90 (2.1%)	8 (1.2%)
UDP-GlcNAc (300)	4134 (100%)	-----
UDP-GlcA (120)	-- --	635 (100%)
UDP-Glc (160)	81 (1.9%)	10 (1.5%)
UDP-GalNAc (280)	74 (1.7%)	19 (2.9%)
UDP-GalA (150)	58 (1.4%)	19 (2.9%)
UDP-GlcNAc + EDTA	31 (0.7%)	-----
UDP-GlcA + EDTA	----	22 (3.4%)

* Membranes (324 ng protein) were incubated at 37°C for 1 h with either 120 μM UDP-[^{14}C] GlcA (2.8x10^{4} dpm) or 300 μM UDP-[^{3}H]GlcNAc (2x10^{4} dpm). The radiolabeled sugar nucleotide was used in the presence of the indicated second nonlabeled sugar nucleotide. HA synthase activity was determined as described in the application.

The term "modified structure" as used herein denotes a hyaluronic acid polymer containing a sugar or derivative not normally found in the naturally occurring HA polysaccharide. The term "modified size distributio refer to the synthesis of hyaluronic acid molecules of a size distribution not normally found with the native enzyme; the engineered size could be much smaller or larger than normal.

Various hyaluronic acid products of differing size have application in the areas of drug delivery and the generation of an

enzyme of altered structure can be combined with a hyaluronic acid of differing size. Applications in angiogenesis and wound healing are potentially large if hyaluronic acid polymers of about 20 monosaccharides can be made in good quantities. Another particular application for small hyaluronic acid oligosaccharides is in the stabilization of recombinant human proteins used for medical purposes. A major problem with such proteins is their clearance from the blood and a short biological half life. One present solution to this problem is to couple a small molecule shield that prevents the protein from being cleared from the circulation too rapidly. Very small molecular weight hyaluronic acid is well suited for this role and would be nonimmunogenic and biocompatible. Larger molecular weight hyaluronic acid attached to a drug or protein may be used to target the reticuloendothelial cell system which has endocytic receptors for hyaluronic acid.

One of ordinary skill in the art given this disclosure would appreciate that there are several ways in which the size distribution of the hyaluronic acid polymer made by the hyaluronate synthase could be regulated to give different sizes. First, the kinetic control of product size can be altered by decreasing temperature, decreasing time of enzyme action and by decreasing the concentration of one or both sugar nucleotide substrates. Decreasing any or all of these variables will give lower amounts and smaller sizes of hyaluronic acid product. The disadvantages of these approaches are that the yield of product will also be

decreased and it may be difficult to achieve reproducibility from day to day or batch to batch.

Secondly, the alteration of the intrinsic ability of the enzyme to synthesize a large hyaluronic acid product. Changes to the protein can be engineered by recombinant DNA technology, including substitution, deletion and addition of specific amino acids (or even the introduction of prosthetic groups through metabolic processing). Such changes that result in an intrinsically slower enzyme could then allow more reproducible control of hyaluronic acid size by kinetic means. The final hyaluronic acid size distribution is determined by certain characteristics of the enzyme, that rely on particular amino acids in the sequence. Among the 20% of residues absolutely conserved between the streptococcal enzymes and the eukaryotic hyaluronate synthases, there is a set of amino acids at unique positions that control or greatly influence the size of the hyaluronic acid polymer that the enzyme can make. Specific changes in any of these residues can produce a modified HAS that produces an HA product having a modified size distribution. Engineered changes to seHAS, spHAS, pmHAS, or cvHAS that decrease the intrinsic size of the hyaluronic acid that the enzyme can make before the hyaluronic acid is released, will provide powerful means to produce hyaluronic acid product of smaller or potentially larger size than the native enzyme.

Finally, larger molecular weight hyaluronic acid made be degraded with specific hyaluronidases to make lower molecular weight hyaluronic acid. This practice, however, is very difficult to

achieve reproducibility and one must meticulously repurify the hyaluronic acid to remove the hyaluronidase and unwanted digestion products.

As shown in FIG. 4, hyaluronan synthase can be engineered to produce hyaluronic acid polymers of different size, in particular smaller, than the normal wildtype enzyme. The figure shows the distribution of HA sizes (in millions of Daltons, a measure of molecular weight) for a series of spHAS enzymes, each of which was engineered by site directed mutagenesis to have a single amino acid change from the native enzyme. Each has a different Cysteine residue replaced with Alanine. The cluster of five curves with open symbols represent the following spHAS proteins: wildtype, C124A, C261A, C366A, and C402A. The filled circles represent the poorly expressed C225A protein which is only partially active.

The filled triangles is the C280A spHAS protein, which is found to synthesize a much smaller range of HA polymers than the normal enzyme or the other variants shown. This reduction to practice shows that it is feasible to engineer the hyaluronate synthase enzyme to synthesize a desired range of HA product sizes. The seHAS, pmHAS, and cvHAS genes encoding hyaluronate synthase can also be manipulated by site directed mutagenesis to produce an enzyme which synthesizes a desired range of HA product sizes.

Structurally modified hyaluronic acid is no different conceptually than altering the size distribution of the hyaluronic acid product by changing particular amino acids in the desired HAS or the spHAS. Derivatives of UDP-GlcNAc, in which the N-acetyl

group is missing UDP-GlcN or replac— with another chemically useful group, are expected to be particula— y useful. The strong substrate specificity must rely on a partic—ar subset of amino acids among the 20% that are conserved. Spec—fic changes to one or more of these residues creates a function— synthase that interacts less specifically with one or more of he substrates than the native enzyme. This altered enzyme could en utilize alternate natural or special sugar nucleotides to incor—rate sugar derivatives designed to allow different chemistries t—be employed for the following purposes: (i) covalently couplin— specific drugs, proteins, or toxins to the structurally modifi— hyaluronic acid for general or targeted drug delivery, radiol—gical procedures, etc. (ii) covalently cross linking the hya—ronic acid itself or to other supports to achieve a gel, or oth · three dimensional biomaterial with stronger physical propertie— and (iii) covalently linking hyaluronic acid to a surface to reate a biocompatible film or monolayer.

Bacteria can also be engine—ed to produce hyaluronic acid. For instance, we have created stra—s of *B. subtilis* containing the spHAS gene, as well as the gene or one of the sugar nucleotide precursors. We chose this bacter since it is frequently used in the biotech industry for the prod—tion of products for human use. These bacteria were intended as fi—st generation prototypes for the generation of a bacterium able to roduce hyaluronic acid in larger amounts than presently available sing a wild type natural strain. We put in multiple copies of thes— genes.

For example, three *Bacillus subtilis* strains were constructed to contain one or both of the *Streptococcus pyogenes* genes for hyaluronan synthase (*spHAS*) and UDP-glucose dehydrogenase, the results of which are shown in Table II-B. Based on a sensitive commercial radiometric assay to detect and quantitate HA, it was determined that the strain with both genes (strain #3) makes and secretes HA into the medium. The parent strain or the strain with just the dehydrogenase gene (strain #1) does not make HA. Strain #2, which contains just the *spHAS* gene alone makes HA, but only 10% of what strain #3 makes. Agarose gel electrophoresis showed that the HA secreted into the medium by strain #3 is very high molecular weight.

TABLE II-B

Strain Number	Cells	Medium(*)	Strain with genes	Cell density (A_{600})
(μg HA per ml of culture)				
1	0	0	hasB	4.8
2	4	35	SpHAS	3.9
3	=>10	>250	SpHAS + hasB	3.2

(*) Most HA is in media but some was cell-associated; HA was determined using the HA Test 50 kit from Pharmacia.

These experiments used the s 'eptococcal promoters normally found with these genes to drive pr=in expression. It is expected that the construction of strains —th the spHAS or seHAS reading frame under control of a *B. subtil*= promoter would yield even more superior results. The vector u=d is a Gram positive/*E. Coli* shuttle vector that has a medium =py number in *B. subtilis* and a gene for erythromycin resistance (=abling resistence to 8 μg/ml in *B. subtilis* or 175 μg/ml in *E. col* . The *B. subtilis* host strain used is 1A1 from BGSC, which ha a tryptophan requirement but otherwise is wildtype, and can =orulate. Cell growth and HA production was in Spizizens Minima=Media plus tryptophan, glucose, trace elements and erthromycin (8 _/ml). Growth was at 32 degrees Celsius with vigorous agitation un 1 the medium was exhausted (~36 hours).

This demonstrates that these ioengineered cells, which would not normally make hyaluronic acid became competent to do so when they are transformed with the spH= gene. The seHAS would also be capable of being introduced into non-hyaluronic acid producing bacteria to create a bioenginee=d bacterial strain capable of producing hyaluronic acid.

A preferred embodiment of th—present invention is a purified composition comprising a polypept le having an amino acid sequence in accordance with SEQ ID NO:2. T= term "purified" as used herein, is intended to refer to an HAS pr=in composition, wherein the HAS protein or appropriately modifie HAS protein (e.g. containing a $[HIS]_6$ tail) is purified to any =gree relative to its naturally-

obtainable state, i.e., in this case, relative to its purity within a prokaryotic cell extract. HAS protein may be isolated from *Streptococcus, Pasturella, chlorella* virus, patient specimens, recombinant cells, infected tissues, isolated subpopulation of tissues that contain high levels of hyaluronate in the extracellular matrix, and the like, as will be known to those of skill in the art, in light of the present disclosure. For instance, the recombinant seHAS or spHAS protein makes up approximately 10% of the total membrane protein of *E. coli*. A purified HAS protein composition therefore also refers to a polypeptide having the amino acid sequence of SEQ ID NO:2, free from the environment in which it may naturally occur (FIG. 5).

Turning to the expression of the seHAS gene whether from genomic DNA, or a cDNA, one may proceed to prepare an expression system for the recombinant preparation of the HAS protein. The engineering of DNA segment(s) for expression in a prokaryotic or eukaryotic system may be performed by techniques generally known to those of skill in recombinant expression.

HAS may be successfully expressed in eukaryotic expression systems, however, the inventors aver that bacterial expression systems can be used for the preparation of HAS for all purposes. It is believed that bacterial expression will ultimately have advantages over eukaryotic expression in terms of ease of use, cost of production, and quantity of material obtained thereby.

The purification of streptococcal hyaluronan synthase (seHAS and spHAS) is shown in Table III and FIG. 6. Fractions from various

stages of the purification scheme were analyzed by SDS-PAGE on a 12.5% gel, which was then stained with Coomassie Brilliant Blue R-250. Lanes: molecular weight markers; 1, whole *E.coli* membranes containing the recombinant seHAS-[illegible]; 2, insoluble fraction after detergent solubilization of membranes; 3, detergent solubilized fraction; 4, flow-through from the Ni-NTA chromatography resin; 5-9, five successive washes of the column (two column volumes each); 10, the eluted pure HA synthase which is a single band.

TABLE II

Step	Total Protein (ug)	Specific Activity (mmol/ug/hr.	Total Activity (mol P-GlcA)	Yield (%)	Purification (-fold)
Membranes	3690	1.0	349	100	1.0
Extract	2128	2.2	725	129	2.2
Affinity Column	39	13	90	14	13.1

It is proposed that transformation of host cells with DNA segments encoding HAS will provide convenient means for obtaining a HAS protein. It is also proposed that cDNA, genomic sequences, and combinations thereof, are suitable for eukaryotic expression, as the host cell will, of course, process the genomic transcripts to yield functional mRNA for translation into protein.

Another embodiment of the present invention is a method of preparing a protein composition comprising growing a recombinant host cell comprising a vector that encodes a protein which includes an amino acid sequence in accordance with SEQ ID NO:2 or functionally similar with conserved or semi-conserved amino acid

changes. The host cell will be grown under conditions permitting nucleic acid expression and protein production followed by recovery of the protein so produced. The production of HAS and ultimately HA, including the host cell, conditions permitting nucleic acid expression, protein production and recovery will be known to those of skill in the art in light of the present disclosure of the seHAS gene, and the seHAS gene protein product HAS, and by the methods described herein.

Preferred hosts for the expression of hyaluronic acid are prokaryotes, such as *S. equisimilis*, and other suitable members of the *Streptococcus* species. However, it is also known that HA may be synthesized by heterologous host cells expressing recombinant HA synthase, such as species members of the *Bacillus*, *Enterococcus*, or even *Escherichia* genus. A most preferred host for expression of the HA synthase of the present invention is a bacteria transformed with the HAS gene of the present invention, such as *Lactococcus* species, *Bacillus subtilis* or *E. coli*.

It is similarly believed that almost any eukaryotic expression system may be utilized for the expression of HAS e.g., baculovirus-based, glutamine synthase-based, dihydrofolate reductase-based systems, SV-40 based, adenovirus-based, cytomegalovirus-based, yeast-based, and the like, could be employed. For expression in this manner, one would position the coding sequences adjacent to and under the control of the promoter. It is understood in the art that to bring a coding sequence under the control of such a promoter, one positions the 5' end of the transcription initiation site of the

transcriptional reading frame of the protein between about 1 and about 50 nucleotides "downstream" of (i.e., 3' of) the chosen promoter. Also, *Saccharomyces cerevisiae* yeast expression vector systems, such as pYES2, will also produce HAS under control of the GAL promoter as shown in FIG. 7. FIG. 7 shows that the spHAS enzyme was produced in recombinant yeast using the pYES2 plasmid. When supplied with UDP-GlcA and UDP-GlcNAc, the enzyme makes high molecular weight HA.

Where eukaryotic expression is contemplated, one will also typically desire to incorporate into the transcriptional unit which includes the HAS gene or DNA, an appropriate polyadenylation site (e.g., 5'-AATAAA-3') if one was not contained within the original cloned segment. Typically, the poly A addition site is placed about 30 to 2000 nucleotides "downstream" of the termination site of the protein at a position prior to transcription termination.

It is contemplated that virtually any of the commonly employed host cells can be used in connection with the expression of HAS in accordance herewith. Examples of preferred cell lines for expressing HAS cDNA of the present invention include cell lines typically employed for eukaryotic expression such as 239, AtT-20, HepG2, VERO, HeLa, CHO, WI 38, BHK, COS-7, RIN and MDCK cell lines. This will generally include the steps of providing a recombinant host bearing the recombinant DNA segment encoding the HAS enzyme and capable of expressing the enzyme; culturing the recombinant host in media under conditions that will allow for transcription of the cloned HAS gene or cDNA and appropriate for the production of the

hyaluronic acid; and separating and purifying the HAS enzyme or the secreted hyaluronic acid from the recombinant host.

Generally, the conditions appropriate for expression of the cloned HAS gene or cDNA will depend upon the promoter, the vector, and the host system that is employed. For example, where one employs the *lac* promoter, one will desire to induce transcription through the inclusion of a material that will stimulate *lac* transcription, such as isopropylthiogalactoside. For example, the cloned seHAS gene of the present invention is expressed as a HIS_6 containing protein in E. coli as shown in FIG. 5. Where other promoters are employed, different materials may be needed to induce or otherwise up-regulate transcription.

FIG. 5 depicts the overexpression of recombinant seHAS and spHAS in *E. coli*. Membrane proteins (5mg per lane) were fractionated by SDS-PAGE using a 10% (w/v) gel under reducing conditions. The gel was stained with Coomassie blue R-250, photographed, scanned, and quantitated using a molecular dynamics personal densitometer (model PDSI P60). The position of HA synthase is marked by the arrow. Lane A is native spHAS (Group A); Lane C is native seHAS; Lane E is recombinant seHAS; Lane P is recombinant spHAS; Lane V is vector alone. Standards used were Bio-rad low Mr and shown in kDa.

In addition to obtaining expression of the synthase, one will preferably desire to provide an environment that is conducive to HA synthesis by including appropriate genes encoding enzymes needed for the biosynthesis of sugar nucleotide precursors, or by using growth

media containing substrates for the precursor-supplying enzymes, such as N-acetylglucosamine or glucosamine ($GlcNAc$ or $GlcNH_2$) and glucose (Glc).

One may further desire to incorporate the gene in a host which is defective in the enzyme hyaluronidase, so that the product synthesized by the enzyme will not be degraded in the medium. Furthermore, a host would be chosen to optimize production of HA. For example, a suitable host would be one that produced large quantities of the sugar nucleotide precursors to support the HAS enzyme and allow it to produce large quantities of HA. Such a host may be found naturally or may be made by a variety of techniques including mutagenesis or recombinant DNA technology. The genes for the sugar nucleotide synthesizing enzymes, particularly the UDP-Glc dehydrogenase required to produce UDP-GlcA, could also be isolated and incorporated in a vector along with the HAS gene or cDNA. A preferred embodiment of the present invention is a host containing these ancillary recombinant gene or cDNAs and the amplification of these gene products thereby allowing for increased production of HA.

The means employed for culturing of the host cell is not believed to be particularly crucial. For useful details, one may wish to refer to the disclosure of U.S. Pat. Nos. 4,517,295; 4,801,539; 4,784,990; or 4,780,414; all incorporated herein by reference. Where a prokaryotic host is employed, such as *S. equisimilis*, one may desire to employ a fermentation of the bacteria under anaerobic conditions in CO_2-enriched broth growth media. This allows for a greater production of HA than under aerobic conditions.

Another consideration is that *Streptococcal* cells grown anaerobically do not produce pyrogenic exotoxins. Appropriate growth conditions can be customized for other prokaryotic hosts, as will be known to those of skill in the art, in light of the present disclosure.

Once the appropriate host has been constructed, and cultured under conditions appropriate for the production of HA, one will desire to separate the HA so produced. Typically, the HA will be secreted or otherwise shed by the recombinant organism into the surrounding media, allowing the ready isolation of HA from the media by known techniques. For example, HA can be separated from the cells and debris by filtering and in combination with separation from the media by precipitation by alcohols such as ethanol. Other precipitation agents include organic solvents such as acetone or quaternary organic ammonium salts such as cetyl pyridinium chloride (CPC).

A preferred technique for isolation of HA is described in U.S. Pat. No. 4,517,295, and which is incorporated herein by reference, in which the organic carboxylic acid, trichloroacetic acid, is added to the bacterial suspension at the end of the fermentation. The trichloroacetic acid causes the bacterial cells to clump and die and facilitates the ease of separating these cells and associated debris from HA, the desired product. The clarified supernatant is concentrated and dialyzed to remove low molecular weight contaminants including the organic acid. The aforementioned procedure utilizes filtration through filter cassettes containing

0.22 μm pore size filters. Diafiltration is continued until the conductivity of the solution decreases to approximately 0.5 megaohms.

The concentrated HA is precipitated by adding an excess of reagent grade ethanol or other organic solvent and the precipitated HA is then dried by washing with ethanol and vacuum dried, lyophilized to remove alcohol. The HA can then be redissolved in a borate buffer, pH 8, and precipitated with CPC or certain other organic ammonium salts such as CETAB, a mixed trimethyl ammonium bromide solution at 4 degree(s) Celsius. The precipitated HA is recovered by coarse filtration, resuspended in 1 M NaCl, diafiltered and concentrated as further described in the above referenced patent. The resultant HA is filter sterilized and ready to be converted to an appropriate salt, dry powder or sterile solution, depending on the desired end use.

A. Typical Genetic Engineering Methods Which May Be Employed

If cells without formidable cell membrane barriers are used as host cells, transfection is carried out by the calcium phosphate precipitation method, well known to those of skill in the art. However, other methods may also be used for introducing DNA into cells such as by nuclear injection, cationic lipids, electroporation, protoplast fusion or by the Biolistic(tm) Bioparticle delivery system developed by DuPont (1989). The advantage of using the DuPont system is a high transformation efficiency. If prokaryotic cells or cells which contain substantial cell wall constructions are used, the preferred method of

transfection is calcium treatment using calcium chloride to induce competence or electroporation.

Construction of suitable vectors containing the desired coding and control sequences employ standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to construct the plasmids required. Cleavage is performed by treating with restriction enzyme (or enzymes) in suitable buffer. In general, about 1 μg plasmid or DNA fragments are used with about 1 unit of enzyme in about 20 μl of buffer solution. Appropriate buffers and substrate amounts for particular restriction enzymes are specified by the manufacturer. Incubation times of about 1 hour at 37° C are workable.

After incubations, protein is removed by extraction with phenol and chloroform, and the nucleic acid is recovered from the aqueous fraction by precipitation with ethanol. If blunt ends are required, the preparation is treated for 15 minutes at 15° C with 10 units of Polymerase I (Klenow), phenol-chloroform extracted, and ethanol precipitated. For ligation approximately equimolar amounts of the desired components, suitably end tailored to provide correct matching are treated with about 10 units T4 DNA ligase per 0.5 μg DNA. When cleaved vectors are used as components, it may be useful to prevent religation of the cleaved vector by pretreatment with bacterial alkaline phosphatase.

For analysis to confirm functional sequences in plasmids constructed, the first step was to amplify the plasmid DNA by cloning into specifically competent *E. coli* SURE cells (Stratagene)

by doing transformation at 30-32°C. Second, the recombinant plasmid is used to transform *E. coli* K5 strain Bi8337-41, which can produce the UDP-GlcA precursor, and successful transformants selected by antibiotic resistance as appropriate. Plasmids from the library of transformants are then screened for bacterial colonies that exhibit HA production. These colonies are picked, amplified and the plasmids purified and analyzed by restriction mapping. The plasmids showing indications of a functional HAS gene are then further characterized by any number of sequence analysis techniques which are known by those of ordinary skill in the art.

B. Source and Host Cell Cultures and Vectors

In general, prokaryotes were used for the initial cloning of DNA sequences and construction of the vectors useful in the invention. It is believed that suitable source may be Gram-positive cells, particularly those derived from the Group C *Streptococcal* strains. Bacteria with a single membrane, but a thick cell wall such as *Staphylococci* and *Streptococci* are Gram-positive. Gram-negative bacteria such as *E. coli* contain two discrete membranes rather than one surrounding the cell. Gram-negative organisms tend to have thinner cell walls. The single membrane of the Gram-positive organisms is analogous to the inner plasma membrane of Gram-negative bacteria. The preferred host cells are *Streptococcus* strains that are mutated to become hyaluronidase negative or otherwise inhibited EP144019, EP266578, EP244757). *Streptococcus* strains that have been particularly useful include *S. equisimilis* and *S. zooepidemicus*.

Prokaryotes may also be used for expression. For the expression of HA synthase in a form most likely to accommodate high molecular weight HA synthesis, one may desire to employ *Streptococcus* species such as *S. equisimilis* or *S. zooepidemicus*. The aforementioned strains, as well as *E. coli* W3110 (F-, lambda-, prototrophic, ATCC No. 273325), bacilli such as *Bacillus subtilis*, or other enterobacteriaceae such as *Serratia marcescens*, could be utilized to generate a "super" HAS containing host.

In general, plasmid vectors containing origins of replication and control sequences which are derived from species compatible with the host cell are used in connection with these hosts. The vector ordinarily carries an origin of replication, as well as marking sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species. pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. A pBR plasmid or a pUC plasmid, or other microbial plasmid or phage must also contain, or be modified to contain, promoters which can be used by the microbial organism for expression of its own proteins.

Those promoters most commonly used in recombinant DNA construction include the *lacZ* promoter, tac promoter, the T7 bacteriophage promoter, and tryptophan (trp) promoter system. While these are the most commonly used, other microbial promoters have been discovered and utilized, and details concerning their nucleotide sequences have been published, enabling a skilled worker

to ligate them functionally with pl smid vectors. Also for use with the present invention one may util_ze integration vectors.

In addition to prokaryotes, e_aryotic microbes, such as yeast cultures may also be used. *Sacc=romyces cerevisiae*, or common baker's yeast is the most co ionly used among eukaryotic microorganisms, although a number=f other strains are commonly available. For expression in *Sacc tromyces*, the plasmid YRp7, for example, is commonly used. This p smid already contains the *trp1* gene which provides a selection ma=er for a mutant strain of yeast lacking the ability to grow witho_ tryptophan, for example, ATCC No. 44076 or PEP4-1. The pres=nce of the *trp1* lesion as a characteristic of the yeast host=cell genome then provides an effective environment for detectin transformation by growth in the absence of tryptophan. Suitabl— promoting sequences in yeast vectors include the promoters for ie galactose utilization genes, the 3-phosphoglycerate kinase or =her glycolytic enzymes, such as enolase, glyceraldehyde-3-phosph_e dehydrogenase, hexokinase, pyruvate decarboxylase, phosphof=ctokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, p■sphoglucose isomerase, and glucokinase.

In constructing suitable exp ssion plasmids, the termination sequences associated with these enes are also ligated into the expression vector 3' of the seq_nce desired to be expressed to provide polyadenylation of the mRNA and termination. Other promoters, which have the addit nal advantage of transcription

controlled by growth conditions are the promoter region for alcohol dehydrogenase 2, cytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, and the aforementioned glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Any plasmid vector containing a yeast-compatible promoter, origin of replication and termination sequences is suitable.

In addition to microorganisms, cultures of cells derived from multicellular organisms may also be used as hosts. In principle, any such cell culture is workable, whether from vertebrate or invertebrate culture. However, interest has been greatest in vertebrate cells, and propagation of vertebrate cells in culture has become a routine procedure in recent years. Examples of such useful host cell lines are VERO and HeLa cells, Chinese hamster ovary (CHO) cell lines, and WI38, BHK, COS, and MDCK cell lines.

For use in mammalian cells, the control functions on the expression vectors are often provided by viral material. For example, commonly used promoters are derived from polyoma, Adenovirus 2, bovine papilloma virus and most frequently Simian Virus 40 (SV40). The early and late promoters of SV40 virus are particularly useful because both are obtained easily from the virus as a fragment which also contains the SV40 viral origin of replication. Smaller or larger SV40 fragments may also be used, provided there is included the approximately 250 bp sequence extending from the *Hind III* site toward the *Bgl I* site located in the viral origin of replication.

Further, it is also possible, nd often desirable, to utilize promoter or control sequences norm—ly associated with the desired gene sequence, provided such contr sequences are compatible with the host cell systems. An origin—f replication may be provided either by construction of the vecto to include an exogenous origin, such as may be derived from SV40 r other viral (e.g., Polyoma, Adeno, BPV) source, or may be provi—ed by the host cell chromosomal replication mechanism. If the ve—or is integrated into the host cell chromosome, the latter mecha—sm is often sufficient.

C. Isolation of a *bona fide* H— synthase gene from a highly encapsulated strain of Group *Streptococcus equisimilis*.

The encoded protein, design—ed seHAS, is 417 amino acids (calculated molecular weight of 4—778 and pI of 9.1) and is the smallest member of the HAS famil— identified thus far (FIG. 2). seHAS also migrates anomalously fa— in SDS-PAGE (M_r~42 kDa) (FIGS. 5 and 8).

FIG. 8 is a graphical represer—ation of a Western Blot analysis of recombinant seHAS using specifi antibodies. Group C (C; lane 1) or Group A (A; lane 4) Streptococc— membranes and *E. coli* membranes (9 mg/lane) containing recombinan seHAS (E; lanes 2, 7, and 9) or spHAS (P; lanes 3, 6, 8, and 10) w e fractionated by reducing SDS-PAGE and electrotransferred t— nitrocellulose. Strips of nitrocellulose were probed and eveloped as described in the application using purified IgG f ctions raised to the following regions of spHAS: central domain—peptide E^{147}-T^{161} (lanes 1-4); C-terminus peptide (lanes 5-6); the omplete protein (lanes 7 and 8);

recombinant central domain (lanes 9 and 10). Nonimmune IgG or membranes from cells transformed with vector alone gave no staining as in lane 5.

The seHAS and spHAS protein (previously identified in U.S. Serial No. 08/899,940) encoding sequences are 72% identical. The deduced protein sequence of seHAS was confirmed by reactivity with a synthetic peptide antibody (FIG. 8). Recombinant seHAS expressed in *E. coli* was recovered in membranes as a major protein (FIG. 5) and synthesized very large molecular weight HA in the presence of UDP-GlcNAc and UDP-GlcA in vitro (FIG. 9).

FIG. 9 shows a kinetic analysis of the HA size distributions produced by seHAS and spHAS. *E. coli* membranes containing equal amounts of seHAS or spHAS protein were incubated at 37°C with 1.35 mM UDP-[^{14}C] GlcA (1.3 x 10^3 dpm/nmol) and 3.0 mM UDP-GlcNAc as described in the application. These substrate concentrations are greater than 15 times the respective Km valves. Samples taken at 0.5, 1.0, and 60 min were treated with SDS and chromatographed over Sephacryl S400 HR. The HA profiles in the fractionation range of the column (fractions 12-24) are normalized to the percent of total HA in each fraction. The values above the arrows in the top panel are the MWs (in millions) of HA determined directly in a separate experiment using a Dawn multiangle laser light scattering instrument (Wyatt Technology Corp.). The size distributions of HA synthesized by seHAS (●,■,▲) and spHAS (○,□,_) at 0.5 min (○,●), 1.0 min (□,■) and 60 min (_,▲) are shown as indicated. Analysis showed that seHAS and spHAS are essentially identical in the size distribution of HA

chains they synthesize (FIG. 9). SeAS is twice as fast as spHAS in its ability to make HA.

C.1 Bacterial strains and vectors

The mucoid group C strain D181; *Streptococcus equisimilis*) was obtained from the Rockfeller University Collection. The *E. coli* host strains Sure and XL1-Blue MRF' ere from Stratagene and strain Top10 F' was from Invitrogen. Unless otherwise noted, Streptococci were grown in THY and *E. coli* strains were grown in LB medium. pKK-223 Expression vector was from Pharmacia, PCR 2.1 cloning vector was from Invitrogen, and predigested λ Zap Express TM *Bam* HI/CIAP Vector was from Stratagene.

C.2 Recombinant DNA and Cloning

High molecular mass Genomic DNA from *Streptococcus equisimilis* isolated by the method of Caparon and Scott (as known by those with ordinary skill in the art) was partially digested with *Sau3AI* to an average size of 2-12 kb. The digested DNA was precipitated with ethanol, washed and ligated to the Bam HI/CIAPλ Zap Express vector. Ligated DNA was packaged into phage with a Packagene™ extract obtained from Promega. The titer of the packaged phage library was checked using XL1-Blue MRF' E. coli as a host.

C.3 Degenerate PCR Amplification

Degenerate oligonucleotides were designed based upon conserved sequences among spHAS (*Streptococcus pyogenes*), DG42 (*Xenopus laevis* HAS; 19) and nodC (a *Rhizobium meliloti* nodulation factor; 20) and were used for PCR amplification with D181 genomic DNA as a template. Amplification conditions were 34 cycles at: 94°C for 1 min, 44°C for

1 min, 72°C for 1.5 min followed by a final extension at 72°C for 10 min. Oligonucleotide *HADRF1*. 5'-GAY MGA YRT YTX ACX AAT TAY GCT ATH GAY TTR GG-3' (SEQ ID NO:20; sense strand) corresponds to the sequence D^{259}RCLTNYAIDL (SEQ ID NO:9; spHAS). Oligonucleotide *HACTR1*, 5'-ACG WGT WCC CCA NTC XGY ATT TTT NAD XGT RCA-3' (SEQ ID NO:21; antisense strand) corresponds to the region C^{404}TIKNTEWGTR (SEQ ID NO:10; spHAS). The degeneracy of bases at some positions are represented by nomenclature adopted by the IUPAC in its codes for degenerate bases listed in Table IV.

TABLE IV
IUPAC Codes - Degenerate Bases

The International Union for Pure and Applied Chemistry (IUPAC) has established a standard single-letter designation for degenerate bases. These are:

B	=	C+G+T
D	=	A+G+T
H	=	A+C+T
K	=	T+G
M	=	A+C
N	=	A+C+G+T
R	=	A+G
S	=	G+C
W	=	A+T
V	=	A+C+G
X	=	a minor bases (specified elsewhere)
Y	=	C+T

These two oligonucleotides gave a 459 bp PCR product, which was separated on an agarose gel and purified using the BIO-101 Geneclean kit. This fragment was then cloned into PCR2.1 vector using TOP 10 F' cells as a host according to the manufacturer's directions. Double stranded plasmid DNA was purified from *E. coli* (Top 10 F') using the QIAfilter Plasmid Midi Kit (Qiagen). Two other degenerate

sense primers were also synthesized *HAVAF1*, 5'-GTN GCT GCT GTW RTX CCW WSX TWT AAY GAR GA-3' (SE ID NO:22, corresponding to the region V[66]AAVIPSYNE (SEQ ID NO: .) of spHAS) and *HAVDF1*, 5'-GTX RWT GAY GGN WSX WSN RAX GAT G GC-3' (SEQ ID NO:23, based on V[100]DDGSSNTD (SEQ ID NO:12) of HAS). Two unique antisense primers were synthesized based on sequence of the 459 bp PCR product. These were: *D181.2*, 5'-A GGA CTT GTT CCA GCG GT-3' (SEQ ID NO:13) and *D181.4*, 5'-TGA G TTC CGA CAC AGG GC-3' (SEQ ID NO:14). Each of the two degenerate sense primers, when used with either D181.2 or D181.4 to amplify D181 genomic DNA, gave expected size PCR products. The four PCR products were cloned and sequenced using the same strategy as above. For each PCR product, sequences obtained from six different clones were compared in order to derive a consensus sequence. Thus we obtained a 1042 bp sequence with a continuous ORF with high homology to spHAS.

C.4 Library Screening

Two molecular probes were used to screen the library; the cloned 459 bp PCR product and oligonucleotide *D181.5* (5'-GCTTGATAGGTCACCAGTGTCACG-3' (SEQ ID NO:15); derived from the 1042 bp sequence). The 459 bp PCR product was radiolabeled using the Prime-It II random primer labeling Kit (Stratagene) according to the manufacturers instructions. Oligonucleotides were labeled by Kinace-It Kinasing Kit (Stratagene) using [γ^{32}P]ATP. Radiolabeled products were separated from nonlabeled material on NucTrap Push columns (Stratagene). The oligoprobe hybridized specifically with a D181 genomic digest on Southern blots. To screen the λ phage library, XLBLUE MRF' was used as a host (3000 plaques/plate) on

Nitrocellulose membranes containing adsorbed phage, were prehybridized at 60°C and hybridized with 5'-end labeled oligonucleotide, D181.5, in QuikHyb Hybridization solution (Stratagene) at 80°C according to instructions.

The membranes were then washed with 2x SSC buffer and 0.1% (w/v) SDS at room temperature for 15 min, at 60°C with 0.1x SSC buffer and 0.1% SDS (w/v) for 30 min, dried and then exposed to Bio-Max MS film overnight at -70°C. Positive plaques were replated and rescreened twice. Pure positive phages were saved in SM buffer with chloroform. PCR on these phages with vector primers revealed 3 different insert sizes.

PCR with a combination of vector primers and primers from different regions of the cloned 1042 bp sequence revealed that only one of the three different phages had the complete HAS gene. The insert size in this phage was 6.5 kb. Attempts to subclone the insert into plasmid form by autoexcision from the selected phage library clone failed. Therefore, a PCR strategy was applied again on the pure positive phage DNA to obtain the 5' and 3' end of the ORF. Oligonucleotide primers *D181.3* (5'-GCCCTGTGTCGGAACATTCA-3' (SEQ ID NO:16)) and T3 (vector primer) amplified a 3kb product and oligonucleotides *D181.5* and T7 (vector primer) amplified a 2.5 kb product. The 5' and 3'-end sequences of the ORF were obtained by sequencing these two above products. Analysis of all PCR product sequences allowed us to reconstruct the ORF of the 1254 bp seHAS gene.

C.5 Expression cloning of the seHAS

Primers were designed at the start and stop codon regions of seHAS to contain an *EcoR1* restriction site in the sense oligonucleotide (5'-AGGATCCGAATTCAGAGAACATTAAAAAACCTC-3' (SEQ ID NO:17)) and a *Pst1* site in the antisense oligonucleotide (5'-AGAATTCTGCAGTTATAATAATTTTTTACGTGT -' (SEQ ID NO:18)). These primers amplified a 1.2 kb PCR product from D181 genomic DNA as well as from pure hybridization-positive phage. The 1.2 kb product was purified by agarose gel electrophoresis, digested with *Pst1* and *EcoR1* and cloned directionally into *Pst1*-and *EcoR1*-digested pKK223 vector. The ligated vector was transformed into *E. coli* SURE cells that were then grown at 30°C. This step was practically important since other host cells or higher temperatures resulted in deletions of the cloned insert. Colonies were isolated and their pDNA purified. Out of six colonies (named a,b,c,d,e, and), five had the correct size insert, while one had no insert.

C.6 HA Synthase Activity

HA synthase activity was assayed in membranes prepared from the 5 above clones. Fresh log phase cells were harvested at 3000g, washed at 4°C with PBS and membranes were isolated by a modification of a protoplast method as known by those of ordinary skill in the art. Membrane preparations from *Streptococcus pyogenes* and *Streptococcus equisimilis* were also obtained by modification of a different protoplast procedure. Membranes were incubated at 37°C in 50 mM sodium and potassium phosphate, pH 7.0 with 20 mM MgCl, 1 mM DTE, 120 μM UDP-GlcA and 300 μM UDP-GlcNAc. Incorporation of sugar

was monitored by using UDP-[^{14}C]GlcA (318 mCi/mmol; ICN) and/or UDP-[^{3}H]GlcNAc (29.2 Ci/mmol NEN). Reactions were terminated by addition of SDS to a final concentration of 2% (w/v). Product HA was separated from precursors by descending paper chromatography and measured by determining incorporated radioactivity at the origin.

C.7 Gel Filtration Analysis

Radiolabeled HA produced *in vitro* by membranes containing recombinant seHAS or spHAS was analyzed by chromatography on a column (0.9 x 40 cm) of Sephacryl S500 HR (Pharmacia Biotech Inc.). Samples (0.4 ml in 200 mM NaCl, 5mM Tris-HCl, pH 8.0, plus 0.5% SDS) were eluted with 200 mM, NaCl, 5 mM Tris-HCL, and pH 8.0 and 0.5 ml fractions were assessed for ^{14}C and/or ^{3}H radioactivity. Authenticity of the HA polysaccharide was assessed by treatment of a separate identical sample with the HA-specific hyaluronate lyase of *Streptomyces hyalurolyticus* (EC 4.2.2.1) at 37°C for 3 hrs. The digest was then subjected to gel filtration.

C.8 SDS-PAGE and Western Blotting

SDS-PAGE was performed according to the Laemmli method. Electrotransfers to nitrocellulose were performed within standard blotting buffer with 20% methanol using a Bio-Rad mini Transblot device. The blots were blocked with 2% BSA in TBS. Protein A/G alkaline phosphatase conjugate (Pierce) and p-nitroblue tetrazolium/5-bromo-4-chloro-3 indolyl phosphate p-toluidine salt were used for detection.

C.9 DNA Sequence and Analysis

Plasmids were sequenced on both strands using fluorescent labeled vector primers. Sequencing reactions were performed using a Thermosequenase™ kit for fluorescent labeled primers (with 7-deazaG). Samples were electrophoresed on a Pharmacia ALF Express DNA Sequencer and data were analyzed by the ALF Manager Software v3.02. Internal regions of inserts were sequenced with internal primers using the ABI Prism 377 (Software version 2.1.1). Ambiguous regions were sequenced manually using Sequenase™ 7-deaza - DNA polymerase, 7-deaza GTP master mix (USB) and [α-^{35}S] dATP (Amersham Life Sciences). The sequences obtained were compiled and analyzed using DNASIS, v2.1 (Hitachi Software Engineering Co., Ltd.). The nucleotide and amino acid sequences were compared with other sequences in the Genbank and other databases.

C.10 Identification of seHAS

Identification of seHAS was accomplished by utilizing a PCR approach with oligonucleotide primers based on several regions of high identity among spHAS, DG42 (not known to be a developmentally regulated *X. laevis* HAS and designated xlHAS) and NodC (a Rhizobium ß-GlcNAc transferase). The xlHAS and NodC proteins are, respectively, ~50% and ~10% identical to spHAS. This strategy yielded a 459 bp PCR product whose sequence was 66.4% identical to spHAS, indicating that a Group C homologue (seHAS) of the Group A (spHAS) HA synthase gene had been identified. The complete coding region of the gene was then reconstructed using a similar PCR-based strategy. A final set of PCR primers was then used to amplify the

complete ORF from genomic DNA. When this 1.2 kb PCR fragment was incorporated into the expression vector pKK223 and transformed into *E. coli* SURE cells, HA synthetic activity was demonstrated in isolated membranes from 5 of the 5 colonies tested.

The ORF of the reconstructed gene encodes a novel predicted protein of 417 amino acids that was not in the database and it is two amino acids shorter than spHAS. The two bacterial proteins are 72% identical and the nucleic acid sequences are 70% identical. The predicted molecular weight of the seHAS protein is 47,778 and the predicted isoelectric point is at pH 9.1. Three recently identified mammalian HASs (muHAS1, muHAS2, muHAS3, FIG. 2) are similar to the bacterial proteins. The overall identity between the two groups is ~28-31%, and in addition many amino acids in seHAS are highly conserved with those of the eukaryotic HASs (e.g. K/R or D/E substitutions). A98R, the PBCV-1 HAS is 28-33 percent identical to the mammalian HASs, and is predicted to have a similar topology in the lipid membrane. Within mammalian species the same family members are almost completely identical (e.g. muHAS1 and huHAS1 are 95% identical; muHAS2 and huHAS2 are 98% identical). However, and as shown in FIG. 3, even within the same species the different HAS family members are more divergent (e.g. muHAS1 and muHAS2 are 53% identical; muHAS1 and muHAS3 are 57% identical; muHAS2 and muHAS3 are 71% identical).

FIG. 10 shows hydropathy plots for seHAS and predicted membrane topology. The hydrophilicity plot for the Streptococcal Group C HAS was generated by the method of Kyte and Doolittle (J. Mol. Biol.

157, 105, 1982) using DNAsis. The protein is predicted to be an integral membrane protein.

FIG. 11 shows a model for the topologic organization of seHAS in the membrane. The proposed topology for the protein conforms to the charge-in rule and puts the large central domain inside. This domain is likely to contain most of the substrate binding and catalytic functions of the enzyme. Cys^{226} in seHAS, which is conserved in all HAS family members as well as the other three cysteines are shown in the central domain. Cys^{281} is a critical residue whose alteration can dramatically alter the size distribution of HA product synthesized by the enzyme.

The overall membrane topology predicted for seHAS is identical to that for spHAS and the eukaryotic HASs reported thus far. The protein has two putative transmembrane domains at the amino terminus and 2-3 membrane-associated or transmembrane domains at the carboxyl end. The hydropathy plots for the two Streptococcal enzymes are virtually identical and illustrate the difficulty in predicting the topology of the extremely hydrophobic region of ~90 residues at K^{313}-R^{406} in seHAS (K^{313}-K^{405} in spHAS).

seHAS was efficiently expressed in *E. coli* cells. Roughly 10% of the total membrane protein was seHAS as assessed by staining of SDS-PAGE gels (FIG. 5). The prominent seHAS band at 42 kD is quantitatively missing in the vector-only control lane. This unusually high level of expression for a membrane protein is also found for spHAS, using the same vector in SURE cells. About 8% of the membrane protein is spHAS in *E. coli* SURE cells. In contrast,

the amount of seHAS in Group C membranes is not more than 1% of the total membrane protein. The spHAS in Group A membranes is barely detectable. The recombinant seHAS expressed in *E. coli* SURE cells does not synthesize HA *in vivo*, since these cells lack UDP-GlcA, one of the required substrates. Membranes, however containing the recombinant seHAS protein synthesize HA when provided with the substrates UDP-GlcNAc and UDP-GlcA (FIG. 12).

FIG. 12 shows the synthesis of authentic HA by recombinant seHAS. *E. coli* membranes (69 μg) prepare from cells containing recombinant seHAS or vector alone were incubated at 37°C for 1 hour with 700 μM UDP-[^{3}H]GlcNAc (2.78 x 10^3 dpm/nmol; □,■) and 300 μM UDP-[^{14}C]GlcA (3.83 x 10^3 dpm/nmol; O,●) in a final volume of 200 μl as described herein. The enzyme reaction was stopped by addition of EDTA to a final concentration of 25 mM. Half the reaction mix was treated with *Streptomyces* hyaluronidase at 37°C for 3 hours. SDS (2%, w/v) was added to hyaluronidase-treated (O,□) and untreated (●,■) samples, which were heated at 90°C for 1 min. The samples were diluted to 500 μl with column buffer (5 mM Tris, 0.2 M Nacl, pH 8.0), clarified by centrifugation and 200 μl was injected onto a Sephacryl S-500 HR column. Fractions (1 ml) were collected and radioactivity was determined. BD is the peak elution position position of blue dextran (~2 x 10^6 DA; Pharmacia). V_o marks the excluded volume and V_i the included volume. The ratio of [^{14}c] GlcA: [^{3}H] GlcNAc incorporated into the total amount of HA fractionated on the column is 1.4, which is identical to the ratio of specific activities of the two substrates. Therefore, the molar ratios of

the sugars incorporated into product is 1:1 as predicted for authentic HA. Membranes from cells transformed with vector alone did not synthesize HA.

Using 120 μM UDP-GlcA and 300 μM UDP-GlcNAc, HA synthesis was linear with membrane protein (at ≤0 μg) and for at least 1 hour. Also, membranes prepared from non-transformed cells or cells transformed with vector alone have no detectable HAS activity. HA synthesis is negligible if Mg^{+2} is chelated with EDTA (<5% of control) or if either of the two substrates are omitted (~2% of control). Recombinant seHAS also showed the expected specificity for sugar nucleotide substrates, being unable to copolymerize either UDP-GalA, UDP-Glc or UDP-GalNAc with either of the two normal substrates (Table II).

Based on gel filtration analysis, the average mass of the HA synthesized by seHAS in isolated membranes is 5-10x10^6 Da. The product of the recombinant seHAS is judged to be authentic HA based on the equimolar incorporation of both sugars and its sensitivity to degradation by the specific *Streptomyces* hyaluronidase (FIG. 12). Although the conditions for total HA synthesis were not optimal (since ~90% of one substrate was incorporated into product), the enzyme produced a broad distribution of HA chain lengths. The peak fraction corresponds to an HA mass of 7.5x10^6 Da which is a polymer containing approximately 36,000 monomeric sugars. The distribution of HA sizes resolved on this column ranged from 2-20x10^6 Da.

The deduced protein sequence of seHAS was confirmed by the ability of antibodies to the spHAS protein to cross-react with the

Group C protein (FIG. 8). Polyclonal antibodies to the whole spHAS protein or to just the central domain of spHAS also reacted with the seHAS protein. Antipeptide antibody to the C-terminus of spHAS did not cross-react with this somewhat divergent region in the seHAS protein. However, antipeptide antibody directed against the spHAS sequence E^{147}-T^{161} recognized the same predicted sequence in seHAS. The antipeptide antibody also reacts with the native seHAS and spHAS proteins in Streptococcal membranes and confirms that the native and recombinant enzymes from both species are of identical size. Like the spHAS protein, seHAS migrates anomalously fast on SDS-PAGE. Although the calculated mass is 47,778 Da, the M_r by SDS-PAGE is consistently ~42 kDa.

Because of the sequence identity within their central domain regions and the overall identical structure predicted for the two bacterial enzymes, the peptide-specific antibody against the region E^{147}-T^{161} can be used to normalize for HAS protein expression in membranes prepared from cells transformed with genes for the two different enzymes. Using this approach, membranes with essentially identical amounts of recombinant spHAS or seHAS were compared with respect to the initial rate of HA synthesis and the distribution of HA product size.

As shown for spHAS, the synthesis of HA chains by seHAS is processive. The enzymes appear to stay associated with a growing HA chain until it is released as a final product. Therefore, it is possible to compare the rates of HA elongation by seHAS and spHAS by monitoring the size distribution of HA chains produced at early

times, during the first round of H chain synthesis. Based on gel filtration analysis of HA product sizes at various times, we estimated that the average rate elongation by seHAS is about 9,000 monosaccharides/minute at 37°C (FIG. 9). In five minutes, the enzymes can polymerize an HA chain of 5-10x10^6 Da. During a 60 min incubation, therefore, each enzyme molecule could potentially initiate, complete and release on the order of 5-8 such large HA molecules. At early times (e.g. ≤ min), reflecting elongation of the first HA chains, the size distribution of HA produced by seHAS was shifted to larger species compared to spHAS. By 60 min the two distributions of HA product sizes are indistinguishable.

The cloned seHAS represents the authentic Group C HA synthase. Previously reported or disclosed "Group C" proteins are, therefore, not the true Group C HAS. The seHAS protein is homologous to nine of the currently known HA synthases from bacteria, vertebrates, and a virus that now comprise this rapidly growing HA synthase family. This homology is shown particularly in FIG. 2. In mammals three genes, designated *HAS 1*, *HAS 2* and *HAS 3*, have been identified and mapped to three different chromosomes in both human and mouse. In amphibians the only HAS protein identified thus far is the developmentally regulated DG42, which was cloned in 1988 and recently shown to encode the HA synthase activity by analysis of the recombinant protein in yeast membranes. Probably other *X. laevus* HAS genes will soon be identified.

A divergent evolution model suggests that a primitive bacterial HAS precursor may have been usurped early during vertebrate

development or the bacterial pathogenic strategy of making an HA capsule was developed when a primitive bacteria captured in primordial HAS. Convergent evolution of the bacterial and eukaryotic HAS enzymes to a common structural solution seems unlikely, but may have occurred.

None of the three mammalian isozymes for HAS have yet been characterized enzymatically with respect to their HA product size. At least ten identified HAS proteins are predicted to be membrane proteins with a similar topology. HA synthesis occurs at the plasma membrane and the HA is either shed into the medium or remains cell associated to form the bacterial capsule or a eukaryotic pericellular coat. The sugar nucleotide substrates in the cytoplasm are utilized to assemble HA chains that are extruded through the membrane to the external space.

The protein topology in the very hydrophobic carboxyl portion of the HAS protein appears to be critical in understanding how the enzymes extend the growing HA chain as it is simultaneously extruded through the membrane. For example, the unprecedented enzymatic activity may require unusual and complex interactions of the protein with the lipid bilayer. Preliminary results based on analysis of spHAS-alkaline phosphatase fusion proteins indicate that the amino and carboxyl termini and the large central domains are all intracellular, as shown in FIGS. 10 and 11. The seHAS protein also contains a large central domain (~63% of the total protein) that appears to contain the two substrate binding sites and the two glycosyltransferase activities needed for HA synthesis. Although

current software programs cannot reliably predict the number or nature of membrane-associated domains within the long C-terminal hydrophobic stretch, the proposed topological arrangement agrees with the present evidence and applies as well to the eukaryotic enzymes, which are ~40% larger primarily due to extention of the C-terminal end of the protein with 2 additional predicted transmembrane domains.

Four of the six Cys residues in spHAS are conserved with seHAs. Only Cys225 in both bacterial enzyme is conserved in all members of the HAS family. Since sulfhydryl reactive agents, such as p-mercurobenzoate or NEM, greatly inhibit HAS activity, it is likely that this conserved Cys is necessary or important for enzyme activity. Initial results from site-directed mutagenesis studies, however, indicate that a C225S mutant of spHAS is not inactive, it retains 5-10% of wildtype activity.

The recognition of nucleic acid sequences encoding only seHAS, only spHAS, or both seHAS and spHAS using specific oligonucleotides is shown in FIG. 13. Three pairs of sense-antisense oligonucleotides were designed based on the sequence of ID SEQ NO. 1 and the coding sequence for spHAS. The seHAS based nucleic acid segments (se1-se2 and sesp1-sesp2) are indicated in FIG. 14. These three oligonucleotide pairs were hybridized under typical PCR reactions with genomic DNA from either Group C (seHAS) (lanes 2, 4, and 6) or Group A (spHAS) (lanes 3, 5, and 7) streptococci. Lanes 1 and 8 indicate the positions of MW standards in kb (kilobases). The PCR reactions were performed using *aq* DNA polymerase (from Promega)

for 25 cycles as follows: 94 degrees Celsius for 1 minute to achieve DNA denaturation, 48 degrees Celsius (42 degrees Celsius for the smaller common sesp primers) for 1 minute to allow hybridization, and 72 degrees Celsius for 1.5 minutes for DNA synthesis. The PCR reaction mixtures were then separated by electrophoresis on a 1% agarose gel.

The se1-se2 primer pair was designed to be uniquely specific for the Group C HAS (seHAS). The sp1-sp2 primer pair was designed to be uniquely specific for the Group A HAS (spHAS). The sesp1-sesp2 primer pair was designed to hybridize to both the Group A and Group C HAS nucleic acid sequences. All three primer pairs behaved as expected, showing the appropriate ability to cross-hybridize and support the generation of PCR products that were specific and/or unique.

The oligonucleotides used for specific PCR or hybridization are shown in FIG. 14. The synthetic oligonucleotides of SEQ ID NOS: 3, 4, 5, and 6 are indicated in the corresponding regions of SEQ ID NO. 1. These regions are in bold face and marked, respectively as primers se1, se2, sesp1, and sesp2. The #1 indicates primers in the sense direction, while the #2 indicates a primer in the antisense direction. Each of the four oligonucleotides will hybridize specifically with the seHAS sequence and the appropriate pairs of sense/antisense primers are suitable for use in the polymerase chain reaction as shown in FIG. 13.

FIG. 7 shows a gel filtration analysis of hyaluronic acid synthesized by recombinant HAS expressed in yeast membranes. A DNA

fragment encoding the open reading frame of 419 amino acid residues corresponding to spHAS (with the original Val codon switched to Met) was subcloned by standard methods in the pYES2 yeast expression vector (from Invitrogen) to produce pYES/HA. Membranes from cells with this construct were prepared by agitation with glass beads. The samples derived from pYES/HA constructs contained substantial HA synthase activity and the "42 kD" HAS protein was detected by Western analysis using specific antibodies; membranes from cells with vector alone possessed neither activity nor the immunoreactive band (not shown). Membranes (315 g protein) were first incubated with carrier free UDP-[^{14}C]GlcA (1 Ci ^{14}C) amd 900 uM unlabeled UDP-GlcNAc in 50 mM Tris, pH 7, 20 mM MgCl2, 1mM DTT, and 0.05 M NaCl (450 ul reaction volume) at 30 degrees Celsius for 1.5 minutes. After this pulse-label period nonradiolabeled UDP-GlcA was then added to final concentrations of 900 uM. Samples (100 uL) were taken after the pulse at 1.5 min (dark circle), and 15 (black square), and 45 (black triangle) min after the "chase." The reactions were terminated by the addition of SDS to 2% and heating at 95 degrees Celsius for 1 min The samples were clarified by centrifugation (10,000 x g, 5 min) before injection of half of the sample onto a Sephacryl S-500HR gel filtration column (Pharmacia; 1 x 50 cm) equilibrated in 0.2 M NaCl, 5 mM Tris, pH 8.

The column was eluted at 0. ml/min and radioactivity in the fractions (1 ml) was quantitated by liquid scintillation counting after adding BioSafeII cocktail .5 ml, Research Products Intl.). The void volume and the totally included volumes were at elution

volumes of 14 ml and 35.5 ml, respectively. The peak of blue dextran (average 2x10 6 Da) eluted at 25-27 ml. The recombinant HAS expressed in the eukaryotic yeast cells makes high molecular weight hyaluronic acid *in vitro*.

Thus it should be apparent that there has been provided in accordance with the present invention a purified nucleic acid segment having a coding region encoding enzymatically active HAS, methods of producing hyaluronic acid from the seHAS gene, and the use of hyaluronic acid produced from a HAS encoded by the seHAS gene, that fully satisfies the objectives and advantages set forth above. Although the invention has been described in conjunction with specific embodiments thereof, it is evident that many alternatives, modifications, and variations will be apparent to those skilled in the art. Accordingly, it is intended to embrace all such alternatives, modifications, and variations that fall within the spirit and broad scope of the appended claims.

Claims

What we claim is:

1. A purified nucleic acid segment comprising a coding region encoding enzymatically active hyaluronate synthase.

2. The purified nucleic acid segment of claim 1, wherein the purified nucleic acid segment encodes the *Streptococcus equisimilis* hyaluronate synthase of SEQ ID NO:2.

3. The purified nucleic acid segment of claim 1, wherein the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

4. A purified nucleic acid segment having a coding region encoding enzymatically active hyaluronate synthase, wherein the purified nucleic acid segment is capable of hybridizing to the nucleotide sequence of SEQ ID NO:1.

5. A purified nucleic acid segment having a coding region encoding enzymatically active hyaluronate synthase, wherein the purified nucleic acid segment has semiconservative or conservative amino codon acid changes when compared to the nucleotide sequence of SEQ ID NO:1.

6. A recombinant vector selected from the group consisting of a plasmid, cosmid, phage, or virus vector and wherein the

recombinant vector further comprises a purified nucleic acid segment having a coding region encoding enzymatically active hyaluronan synthase.

7. The recombinant vector of claim 5, wherein the purified nucleic acid segment encodes the *Streptococcus equisimilis* hyaluronan synthase of SEQ ID NO:2.

8. The recombinant vector of claim 6, wherein the purified nucleic acid segment comprises a nuclectide sequence in accordance with SEQ ID NO:1.

9. The recombinant vector of claim 6, wherein the plasmid further comprises an expression vector.

10. The recombinant vector of claim 9, wherein the expression vector comprises a promoter operatively linked to the enzymatically active *Streptococcus equisimilis* hyaluronan synthase coding region.

11. A recombinant host cell, wherein the recombinant host cell is a prokaryotic cell transformed with a recombinant vector comprising a purified nucleic acid segment having a coding region encoding enzymatically active hyaluronan synthase.

12. The recombinant host cell of claim 11, wherein the purified nucleic acid segment encodes the *Streptococcus equisimilis* hyaluronan synthase of SEQ ID NO:2.

13. The recombinant host cell of claim 11, wherein the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

14. The recombinant host cell of claim 13, wherein the host cell produces hyaluronic acid.

15. The recombinant host cell of claim 11, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified structure.

16. The recombinant host cell of claim 11, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified size distribution.

17. A recombinant host cell, wherein the recombinant host cell is a eukaryotic cell transfected with a recombinant vector comprising a purified nucleic acid segment having a coding region encoding enzymatically active hyaluronan synthase.

18. The recombinant host cell of claim 17, wherein the purified nucleic acid segment encodes the *Streptococcus equisimilis* hyaluronan synthase of SEQ ID NO:2.

19. The recombinant host cell of claim 17, wherein the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

20. The recombinant host cell of claim 19, wherein the host cell produces hyaluronic acid.

21. The recombinant host cell of claim 17, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified structure.

22. The recombinant host cell of claim 17, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified size distribution.

23. A recombinant host cell, wherein the recombinant host cell is electroporated to introduce a recombinant vector into the recombinant host cell, wherein the recombinant vector comprises a purified nucleic acid segment having a coding region encoding enzymatically active hyaluronan synthase.

24. The recombinant host cell of claim 23, wherein the purified nucleic acid segment encodes the *Streptococcus equisimilis* hyaluronan synthase of SEQ ID NO:2.

25. The recombinant host cell of claim 23, wherein the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

26. The recombinant host cell of claim 25, wherein the host cell produces hyaluronic acid.

27. The recombinant host cell of claim 23, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified structure.

28. The recombinant host cell of claim 23, wherein the enzymatically active *Streptococcus equisimilis* hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified size distribution.

29. A recombinant host cell, wherein the recombinant host cell is transduced with a recombinant vector comprising a purified nucleic acid segment having a coding region encoding enzymatically active *Streptococcus equisimilis* hyaluronan synthase.

30. The recombinant host cell of claim 29, wherein the purified nucleic acid segment encodes the *Streptococcus equisimilis* hyaluronan synthase of SEQ ID NO:2.

31. The recombinant host cell of claim 29, wherein the purified nucleic acid segment comprises a nucleotide sequence in accordance with SEQ ID NO:1.

32. The recombinant host cell of claim 31, wherein the host cell produces hyaluronic acid.

33. The recombinant host cell of claim 29, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified structure.

34. The recombinant host cell of claim 29, wherein the enzymatically active hyaluronan synthase is capable of producing a hyaluronic acid polymer having a modified size distribution.

35. A purified composition, wherein the purified composition comprises an enzymatically active hyaluronan synthase polypeptide.

36. A purified composition, wherein the purified composition comprises a polypeptide having an amino acid sequence in accordance with SEQ ID NO:2.

37. A method for detecting a DNA species, comprising the steps of:

obtaining a DNA sample;

contacting the DNA sample with purified nucleic acid segment in accordance with SEQ ID NO:1;

hybridizing the DNA sample and the purified nucleic acid segment thereby forming a hybridized complex; and

detecting the complex.

38. A method for detecting a bacterial cell that expresses mRNA encoding *Streptococcus equisimilis* hyaluronan synthase, comprising the steps of:

obtaining a bacterial cell sample;

contacting at least one nucleic acid from the bacterial cell sample with purified nucleic acid segment in accordance with SEQ ID NO:1;

hybridizing the at least one nucleic acid and the purified nucleic acid segment thereby forming a hybridized complex; and

detecting the hybridized complex, wherein the presence of the hybridized complex is indicative of a bacterial strain that expresses mRNA encoding *Streptococcus equisimilis* hyaluronan synthase.

39. A method for producing hyaluronic acid, comprising the steps of:

introducing a purified nucleic acid segment having a coding region encoding enzymatically active hyaluronan synthase into a host organism, wherein the host organism contains nucleic acid segments encoding enzymes which produce UDP-GlcNAc and UDP-GlcA;

growing the host organism in a medium to secrete hyaluronic acid; and

recovering the secreted hyaluronic acid.

40. The method according to claim 39, wherein the step of recovering the hyaluronic acid comprises extracting the secreted hyaluronic acid from the medium.

41. The method according to claim 40, further comprising the step of purifying the extracted hyaluronic acid.

42. The method according to claim 39, wherein in the step of growing the host organism, the host organism secretes a structurally modified hyaluronic acid.

43. The method according to claim 39, wherein in the step of growing the host organism, the host organism secretes a hyaluronic acid having a modified size.

44. A pharmaceutical composition comprising a preselected pharmaceutical drug and an effective amount of hyaluronic acid produced by hyaluronan synthase.

45. The pharmaceutical composition of claim 44, wherein the hyaluronic acid is produced by the *Streptococcus equisimilis* hyaluronan synthase of SEQ ID NO:2.

46. The pharmaceutical composition according to claim 44, wherein the molecular weight of the hyaluronic acid is modified thereby producing a modified molecular weight pharmaceutical composition capable of evading an immune response.

47. The pharmaceutical composition according to claim 44, wherein the molecular weight of the hyaluronic acid is modified thereby producing a modified molecular weight pharmaceutical composition capable of targeting a specific tissue or cell type within the patient having an affinity for the modified molecular weight pharmaceutical composition.

48. A purified and isolated nucleic acid sequence encoding enzymatically active hyaluronan synthase, the nucleic acid sequence selected from the group consisting of:

(a) the nucleic acid sequence in accordance with SEQ ID NO:1;

(b) complementary nucleic acid sequences to the nucleic acid sequence in accordance with SEQ ID NO:1;

(c) nucleic acid sequences which will hybridize to the nucleic acid in accordance with SEQ ID NO:1;

(d) nucleic acid sequences which will hybridize to the complementary nucleic acid sequences of SEQ ID NO:1; and

(e) nucleic acid sequences which will hybridize to PCR probes selected from the group consisting of PCR probes of SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6.

49. A purified and isolated nucleic acid segment consisting essentially of a nucleic acid segment encoding enzymatically active hyaluronan synthase.

50. A procaryotic or eucaryotic host cell transformed or transfected with an isolated nucleic acid segment according to claim 1, 2, or 3 in a manner allowing the host cell to express hyaluronic acid.

51. An isolated nucleic acid segment consisting essentially of a nucleic acid segment encoding hyaluronan synthase having a nucleic acid segment sufficiently duplicative of the nucleic acid segment in accordance of SEQ ID NO:1 to allow possession of the biological property of encoding for *Streptococcus equisimilis* hyaluronan synthase.

52. A cDNA sequence according to claim 51.

53. A procaryotic or eucaryotic host cell transformed or transfected with a nucleic acid segment according to claim 51 in a manner allowing the host cell to express hyaluronic acid.

54. A purified nucleic acid segment having a coding region encoding enzymatically active hyaluronan synthase, wherein the purified nucleic acid segment is capable of hybridizing to the nucleotide sequence in accordance with SEQ ID NO:1.

55. A purified nucleic acid segment according to SEQ ID NO:3 capable of hybridizing to SEQ ID NO:1.

56. A purified nucleic acid segment according to SEQ ID NO:4 capable of hybridizing to SEQ ID NO:1.

57. A purified nucleic acid segment according to SEQ ID NO:5 capable of hybridizing to SEQ ID NO:1.

58. A purified nucleic acid segment according to SEQ ID NO:6 capable of hybridizing to SEQ ID NO:1.

59. A purified nucleic acid segment having a coding region encoding enzymatically active hyaluronate synthase, the purified nucleic acid segment selected from the group consisting of:

(A) the nucleic acid segment according to SEQ ID NO: 2;

(B) the nucleotide sequence in accordance with SEQ ID NO: 1;

(C) nucleic acid segments which hybridize to the nucleic acid segments defined in (A) or (B) or fragments thereof; and

(D) nucleic acid segments which but for the degeneracy of the genetic code, or encoding of functionally equivalent amino acids, would hybridize to the nucleic acid segments defined in (A), (B), and (C).

60. A purified nucleic acid segment comprising a coding region encoding hyaluronate synthase.

61. A recombinant vector wherein the recombinant vector further comprises a purified nucleic acid segment having a coding region encoding hyaluronan synthase.

62. A recombinant host cell, wherein the host cell is transformed with a recombinant vector comprising a purified nucleic acid segment having a coding region encoding hyaluronan synthase.

63. A recombinant host cell, wherein the recombinant host cell is transfected with a recombinant vector comprising a purified nucleic acid segment having a coding region encoding hyaluronan synthase.

64. A recombinant host cell, wherein the recombinant host cell includes a recombinant vector, wherein the recombinant vector comprises a purified nucleic acid segment having a coding region encoding hyaluronan synthase.

65. A purified composition, wherein the purified composition comprises a hyaluronan synthase polypeptide.

66. A method for detecting a DNA species, comprising the steps of:

contacting a DNA sample with a purified nucleic acid segment,

hybridizing the DNA sample and the purified nucleic acid segment thereby forming a hybridized complex;

and detecting the complex.

67. A method for detecting a bacterial cell that expresses mRNA encoding *Streptococcus equisimilis* hyaluronan synthase, comprising the steps of:

contacting at least one nucleic acid from a bacterial cell sample with a purified nucleic acid segment;

hybridizing the at least one nucleic acid and the purified nucleic acid segment

thereby forming a hybridized complex; and

detecting the hybridized complex.

68. A method for producing hyaluronic acid, comprising the steps of:

introducing a purified nucleic acid segment having a coding region encoding hyaluronan synthase into a host organism;

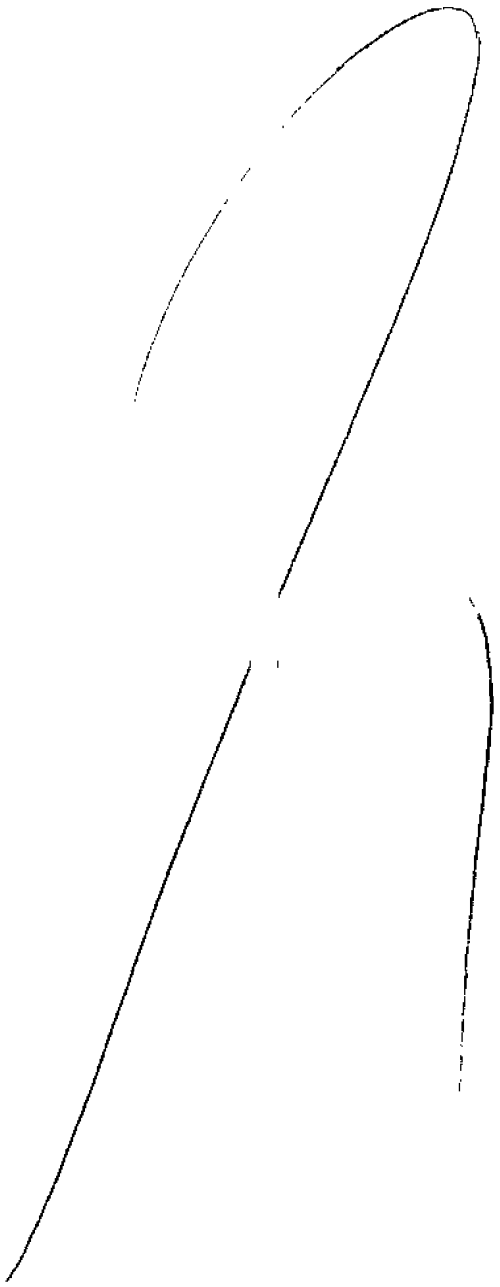
growing the host organism to secrete hyaluronic acid; and

recovering the secreted hyaluronic acid.

69. A purified and isolated nucleic acid segment comprising a nucleic acid segment encoding hyaluronan synthase.

70. An isolated nucleic acid segment comprising a nucleic acid segment encoding hyaluronan synthase.

71. A cDNA sequence according to claim 70.



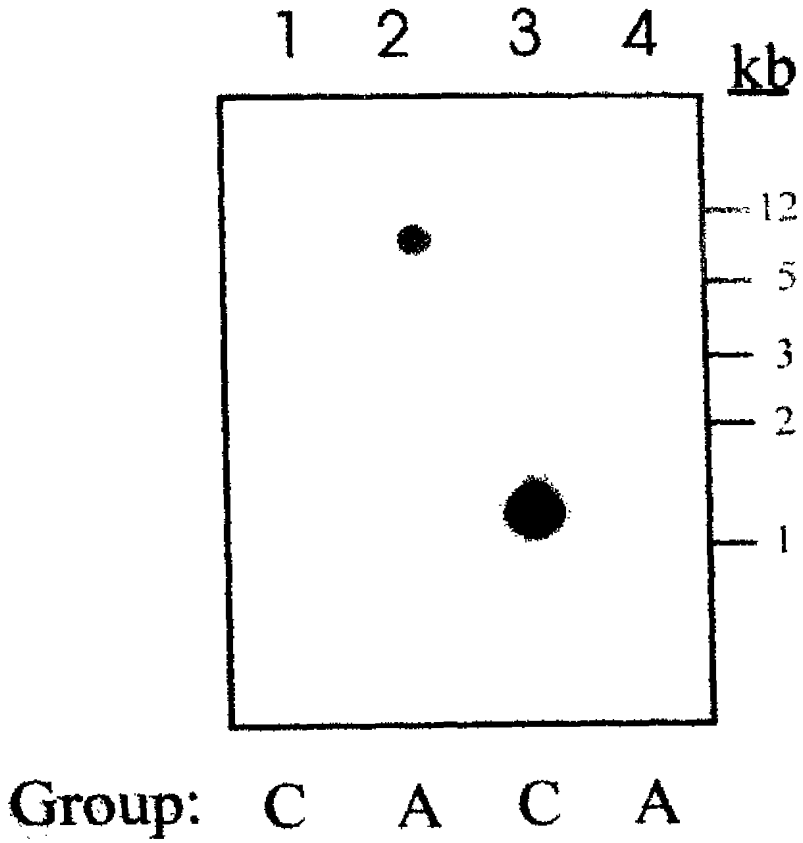
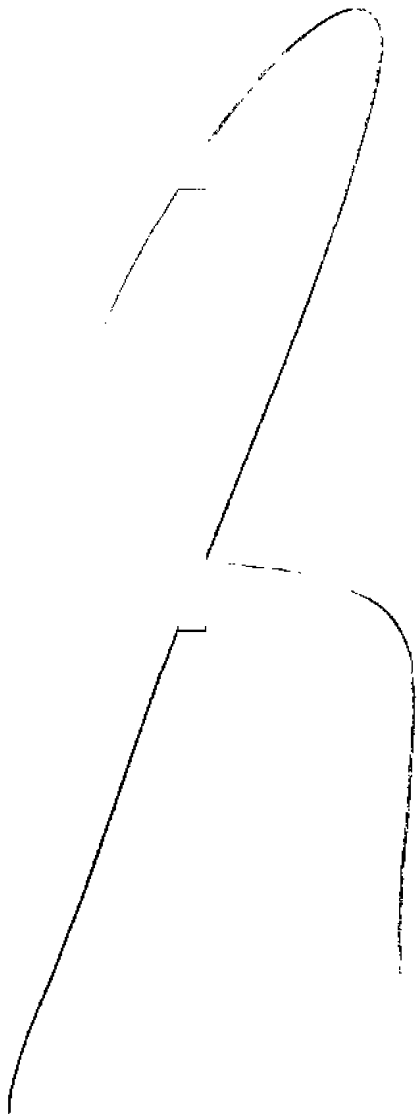


Fig. 1



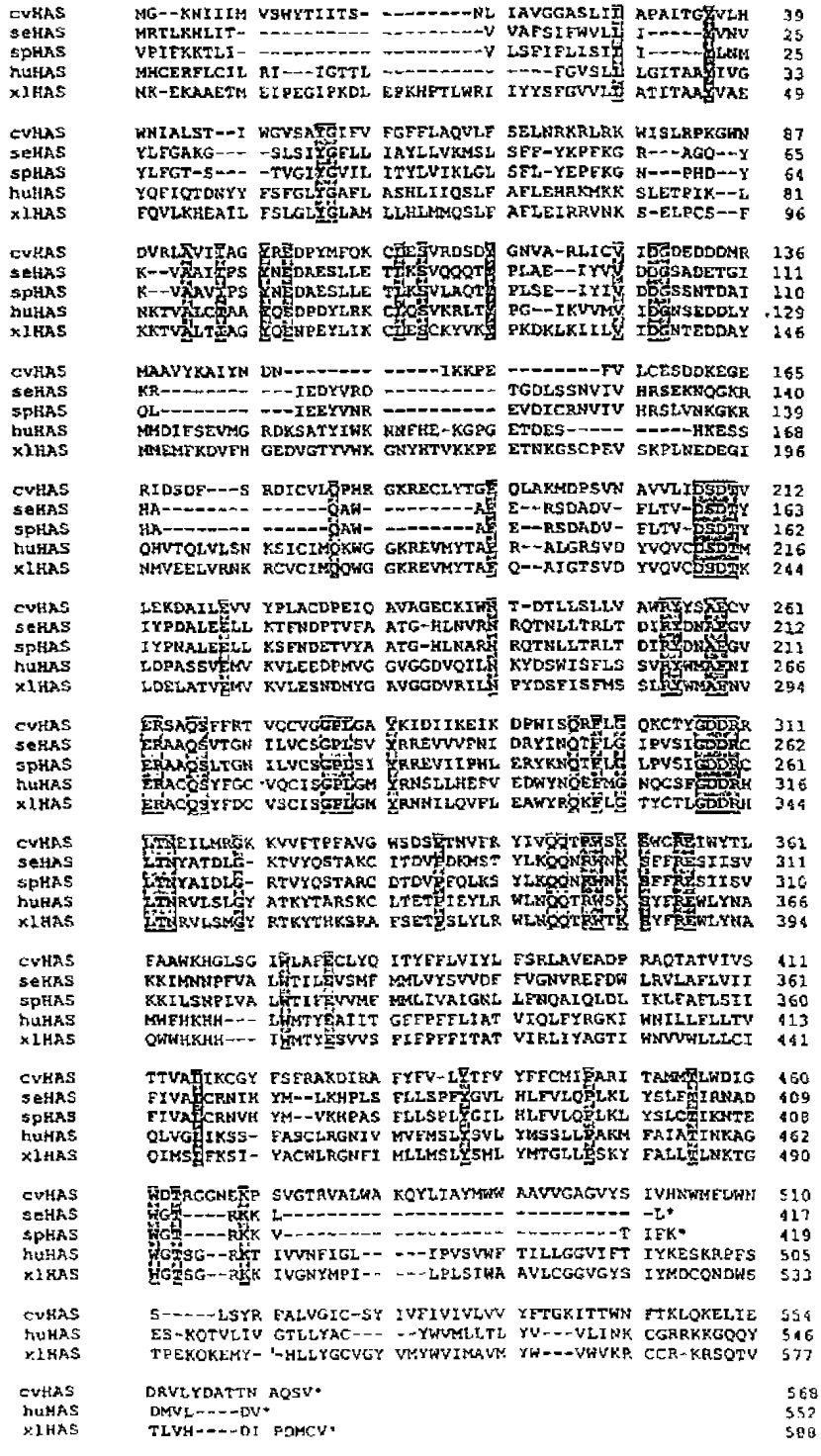
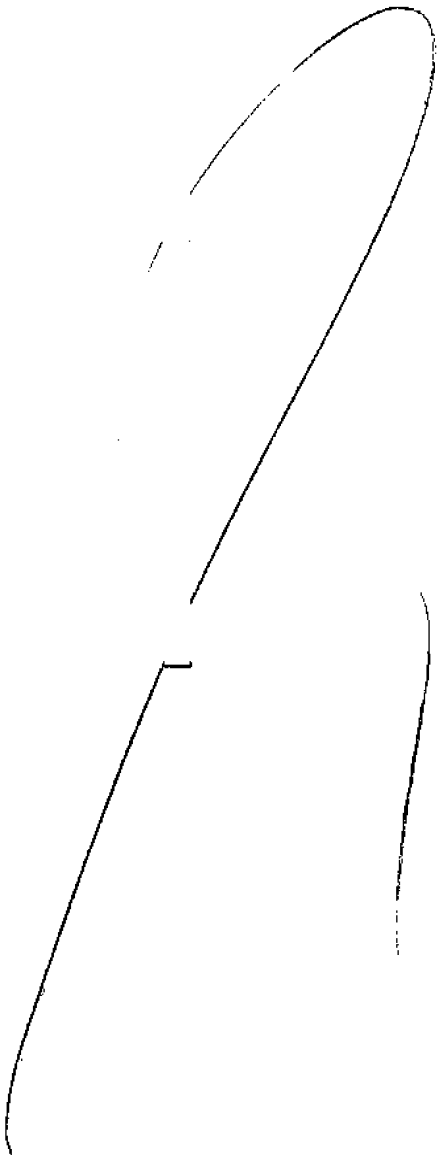


FIG. 2



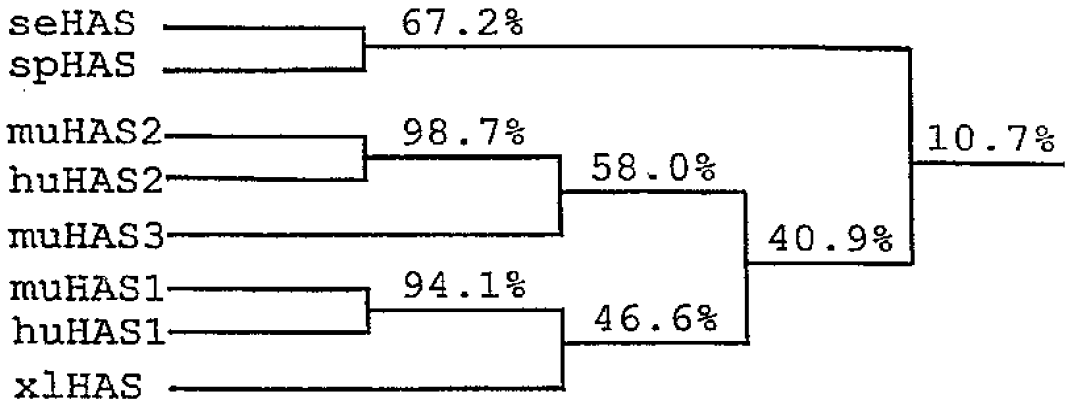
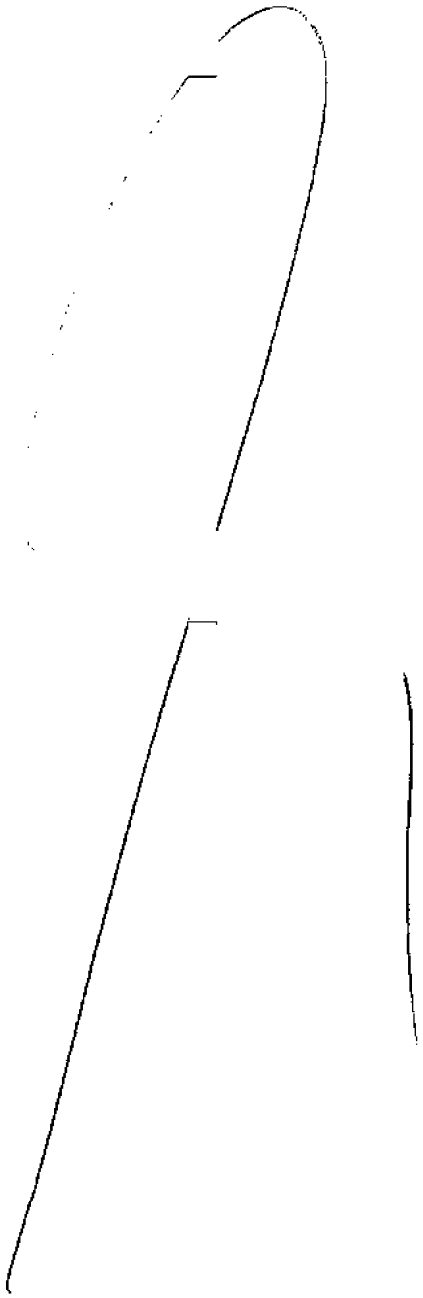

seHAS
spHAS
muHAS2
huHAS2
muHAS3
muHAS1
huHAS1
xlHAS
67.2%
98.7%
58.0%
94.1%
46.6%
40.9%
10.7%

FIG. 3



SIZE DISTRIBUTION OF HYALURONAN PRODUCED BY DIFFERENT ENGINEERED STREPTOCOCCAL HAS ENZYMES

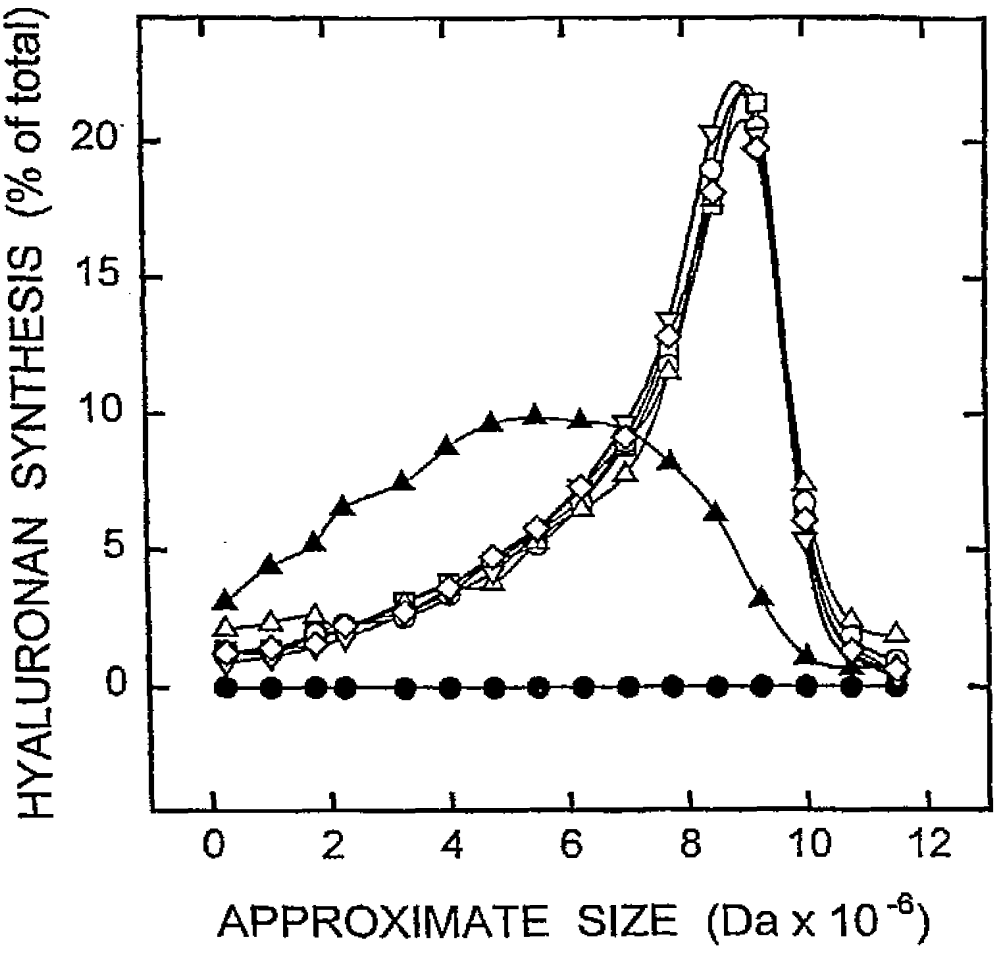
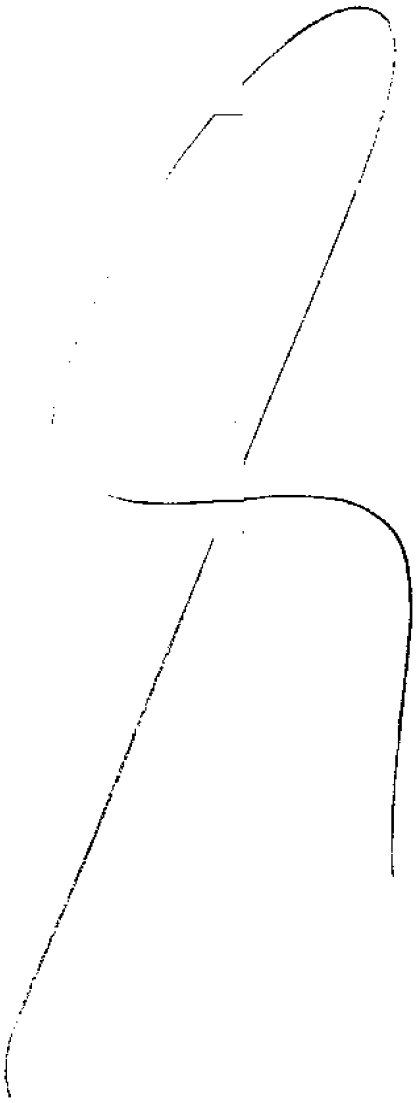


FIG. 4



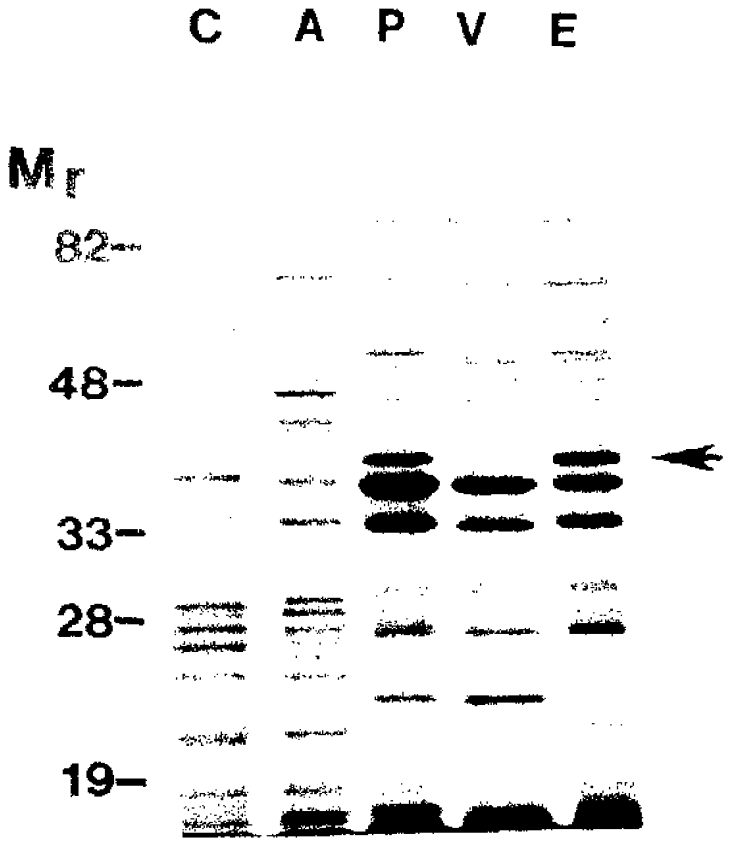
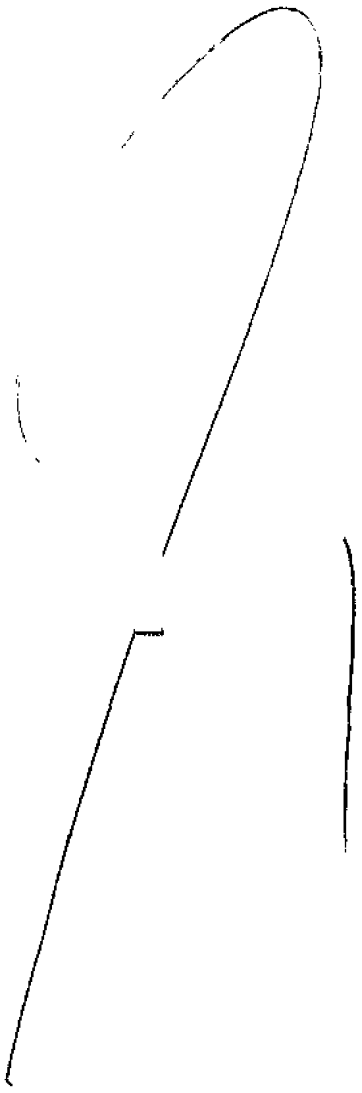


Fig. 5



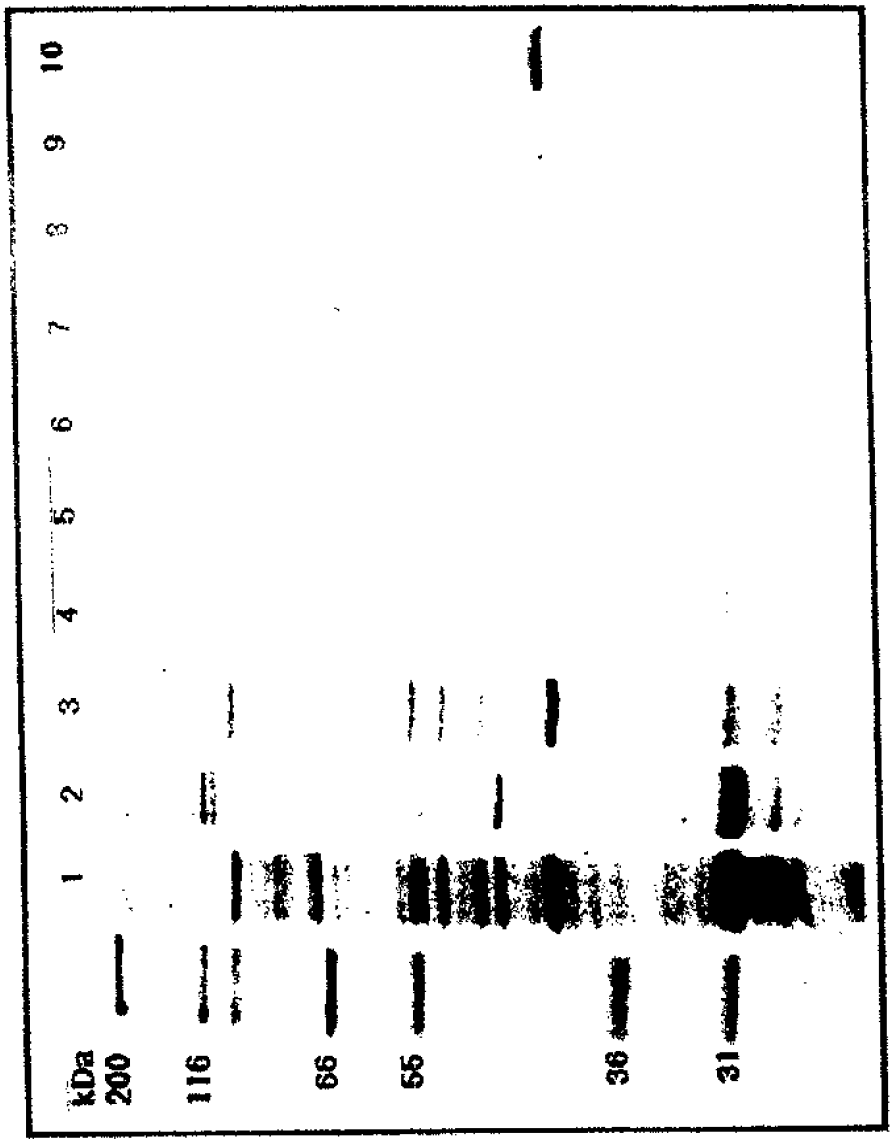
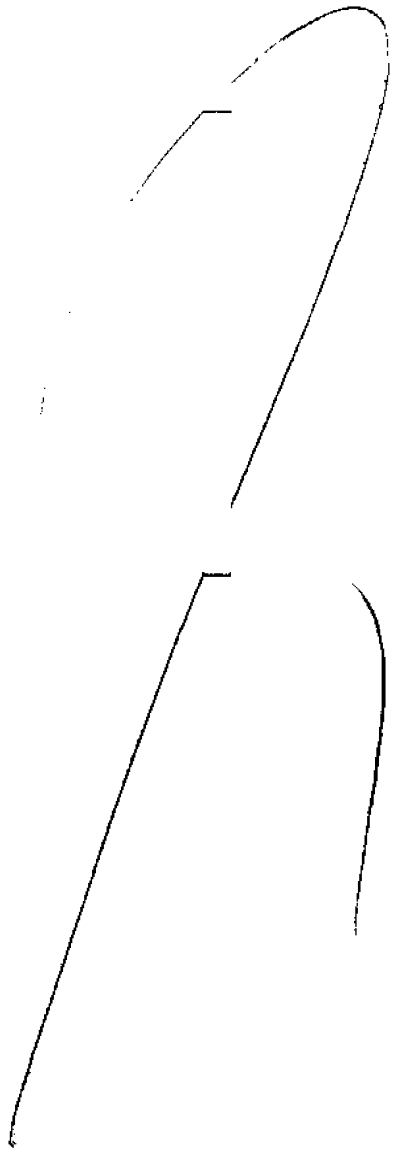


Fig. 6



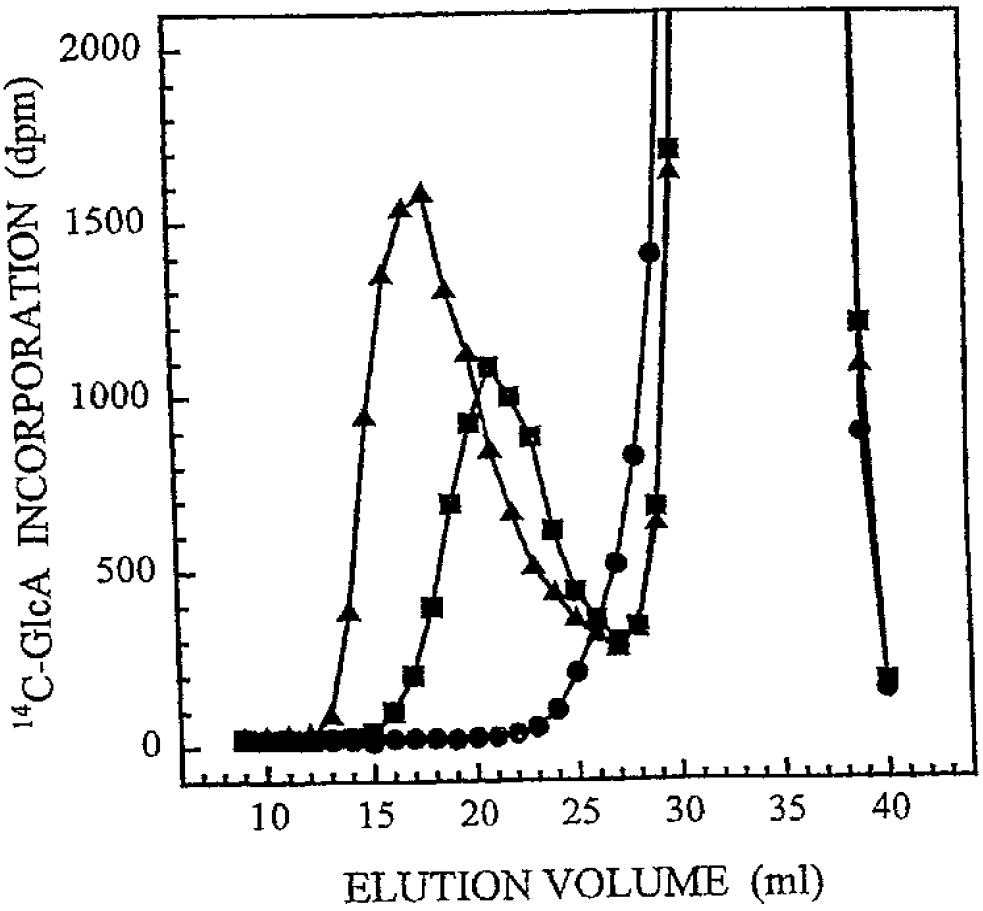
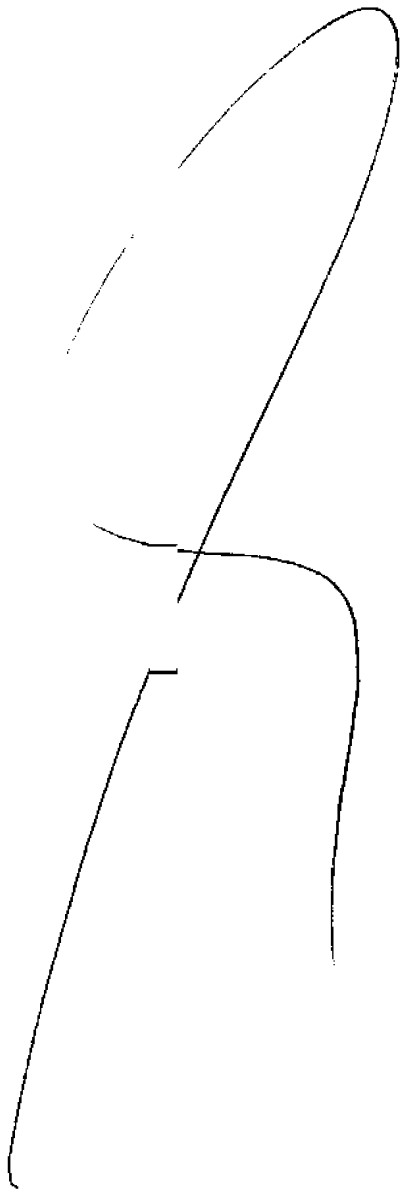

2000
1500
1000
500
0
^{14}C-GlcA INCORPORATION (dpm)
10
15
20
25
30
35
40
ELUTION VOLUME (ml)

FIG. 7



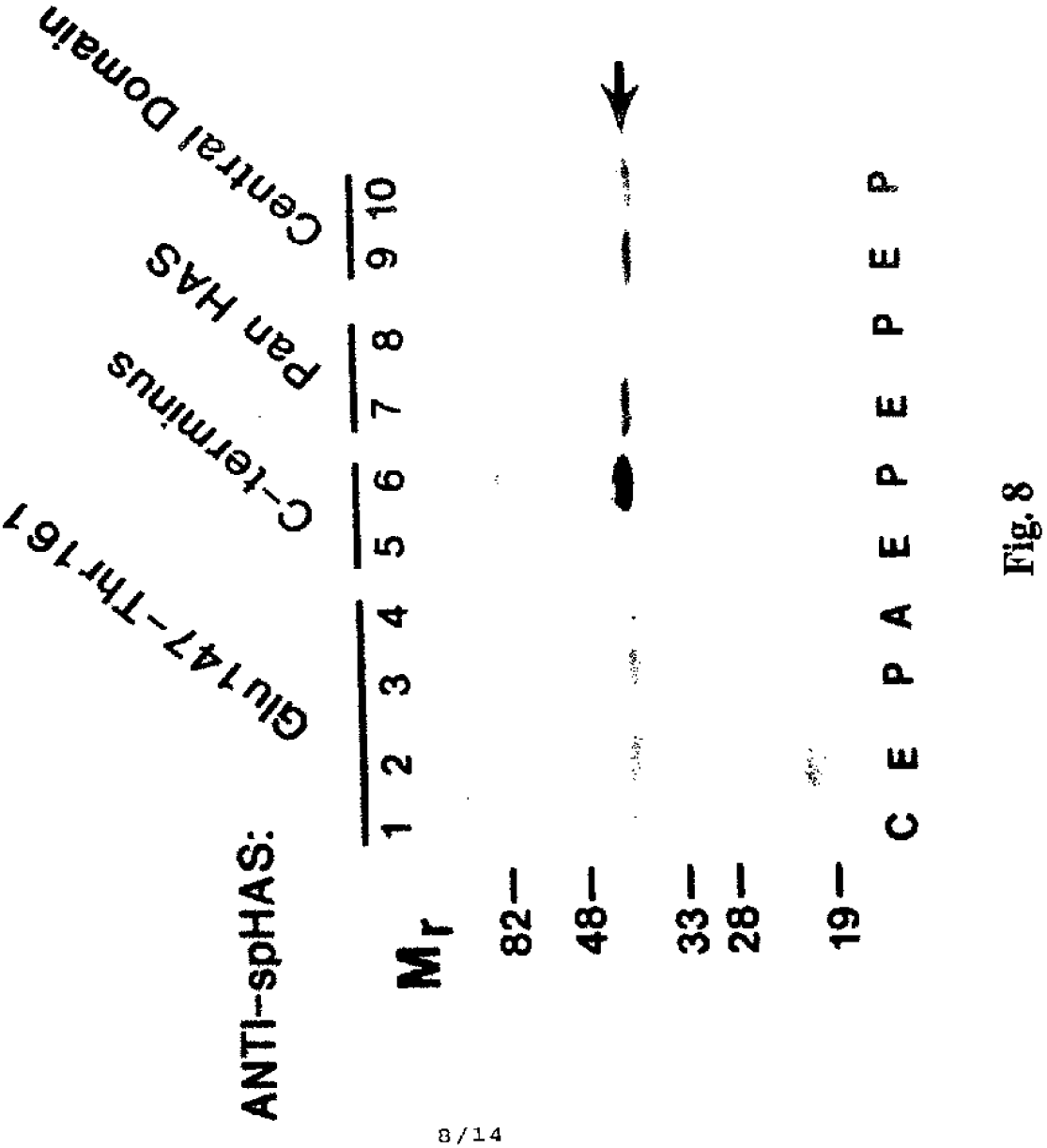
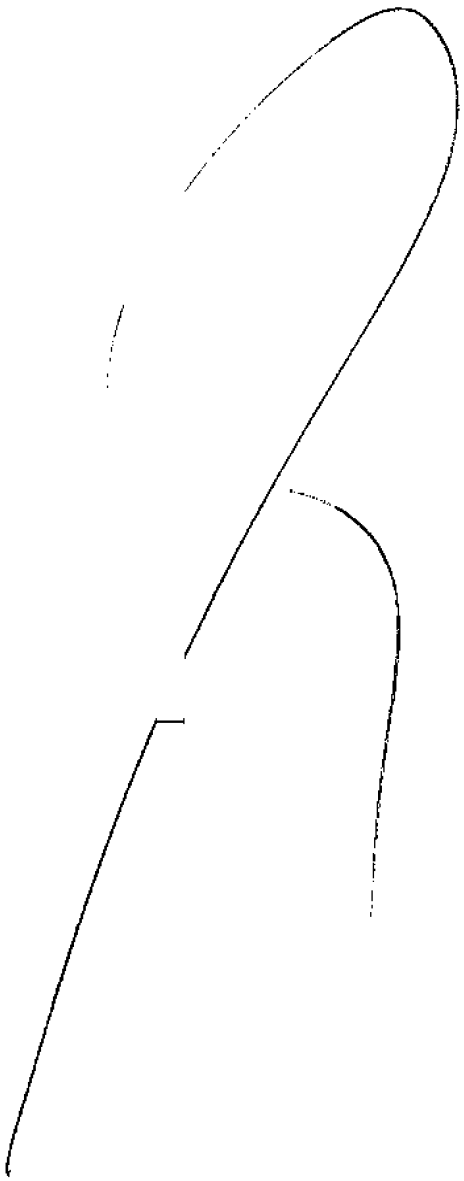

ANTI-spHAS:
Glu147-Thr161
C-terminus
Pan HAS
Central Domain
1 2 3 4 5 6 7 8 9 10
Mr
82
48
33
28
19
C E P A E P E P E P

Fig. 8



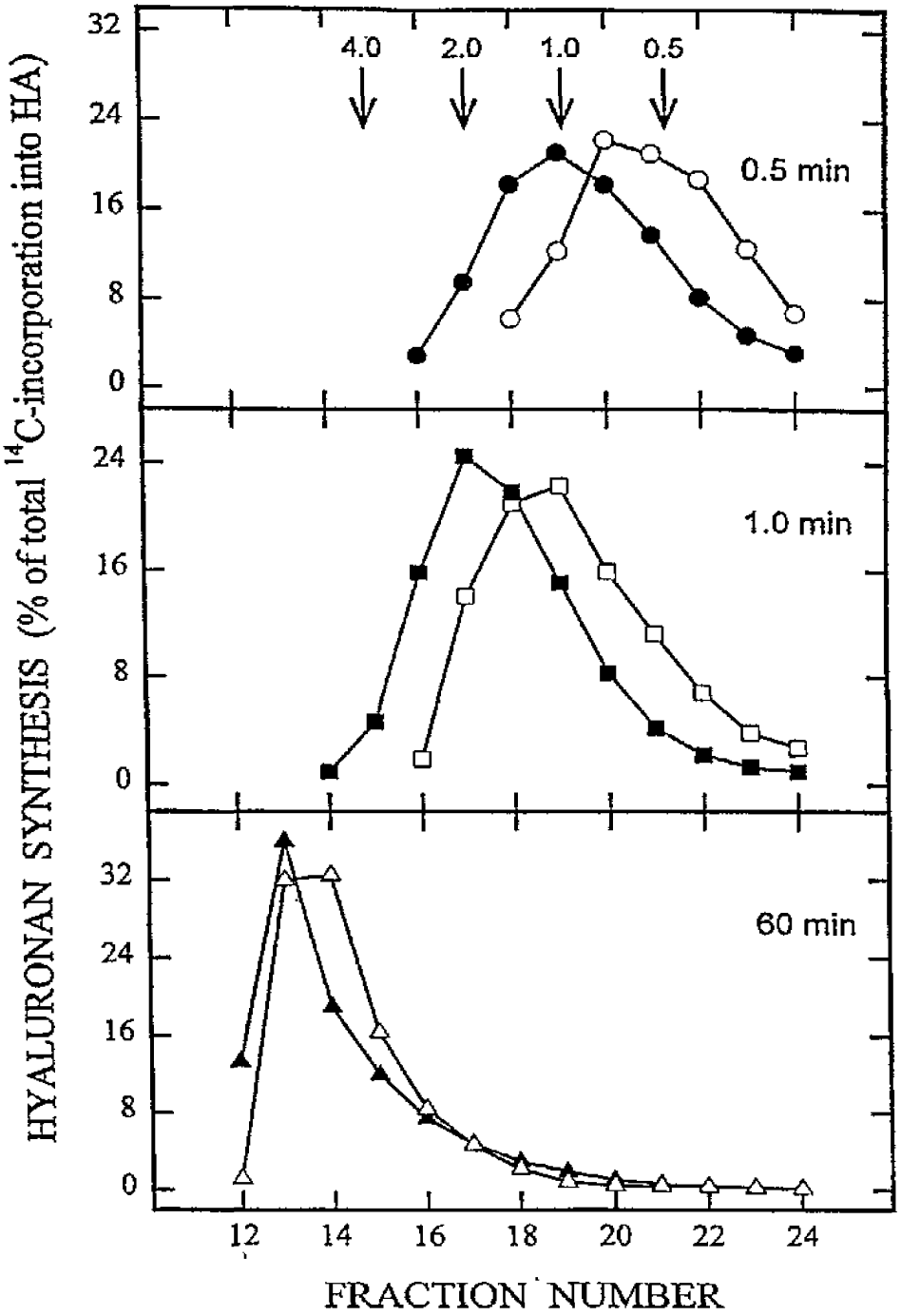
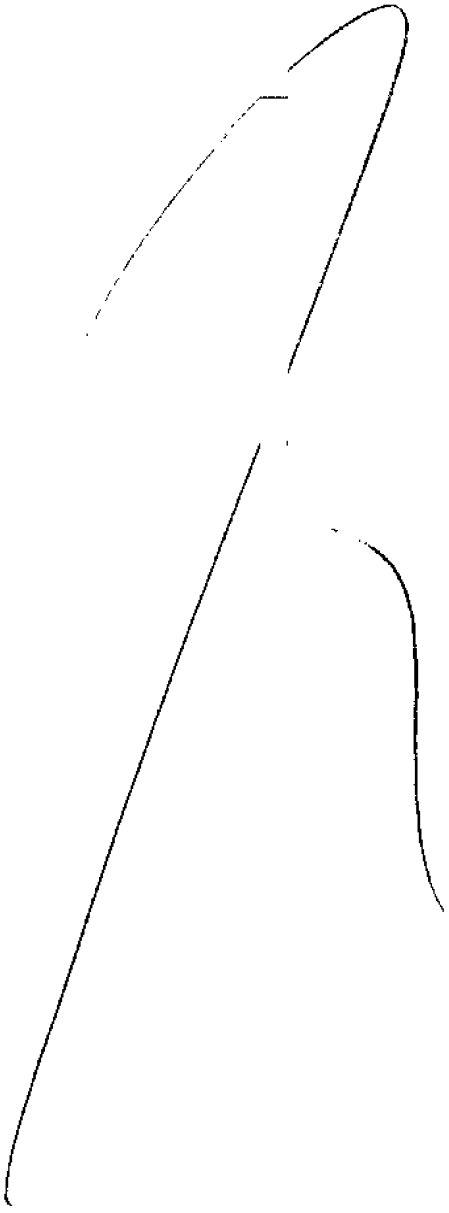

4.0
2.0
1.0
0.5
0.5 min
1.0 min
60 min
32
24
16
8
0
12
14
16
18
20
22
24
HYALURONAN SYNTHESIS (% of total ^{14}C-incorporation into HA)
FRACTION NUMBER

FIG. 9



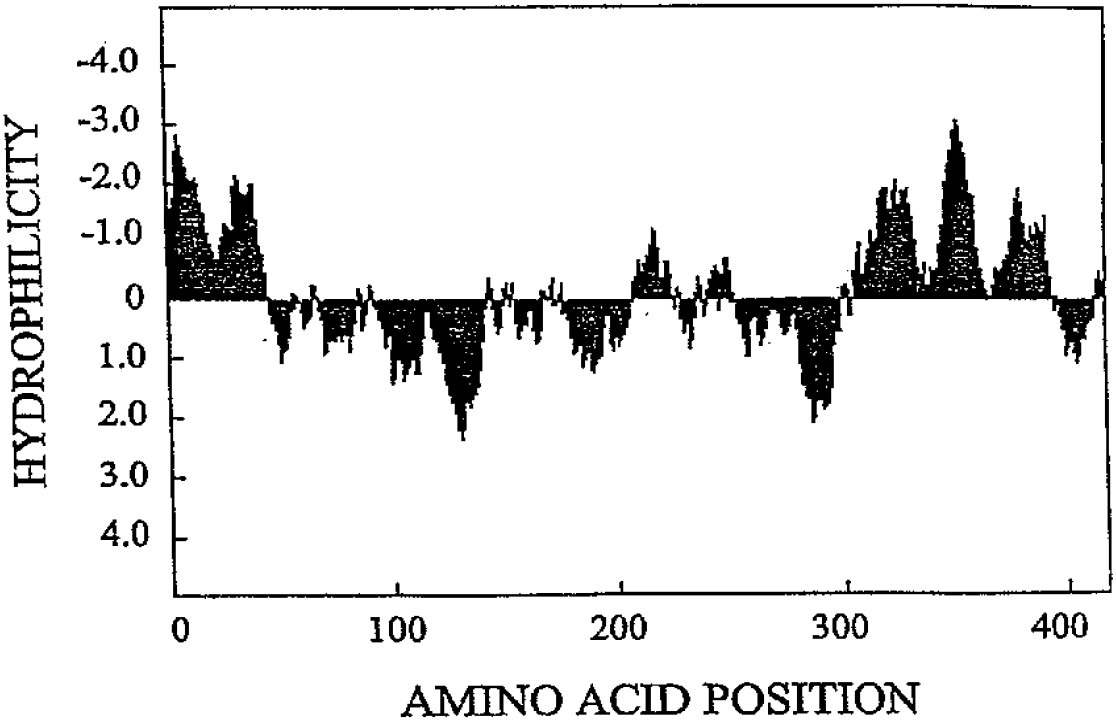
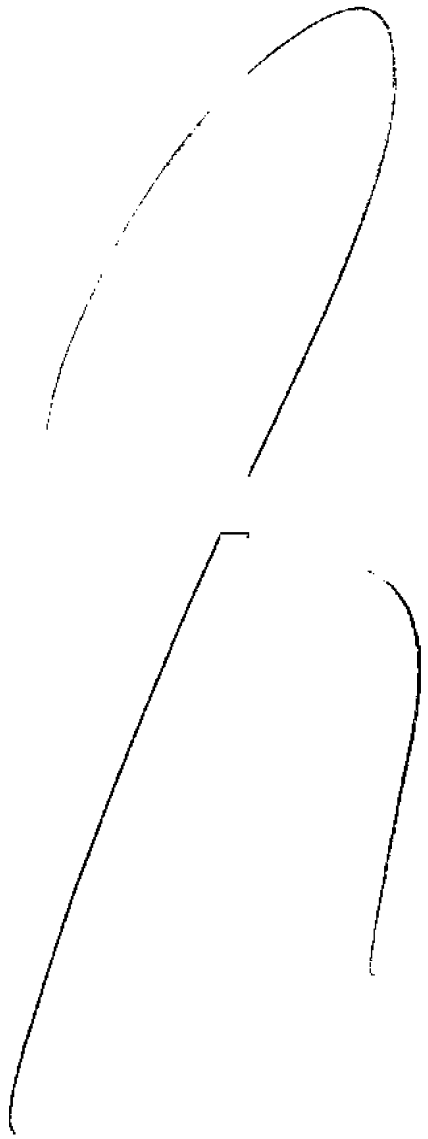

HYDROPHILICITY
-4.0
-3.0
-2.0
-1.0
0
1.0
2.0
3.0
4.0
0
100
200
300
400
AMINO ACID POSITION

FIG. 10



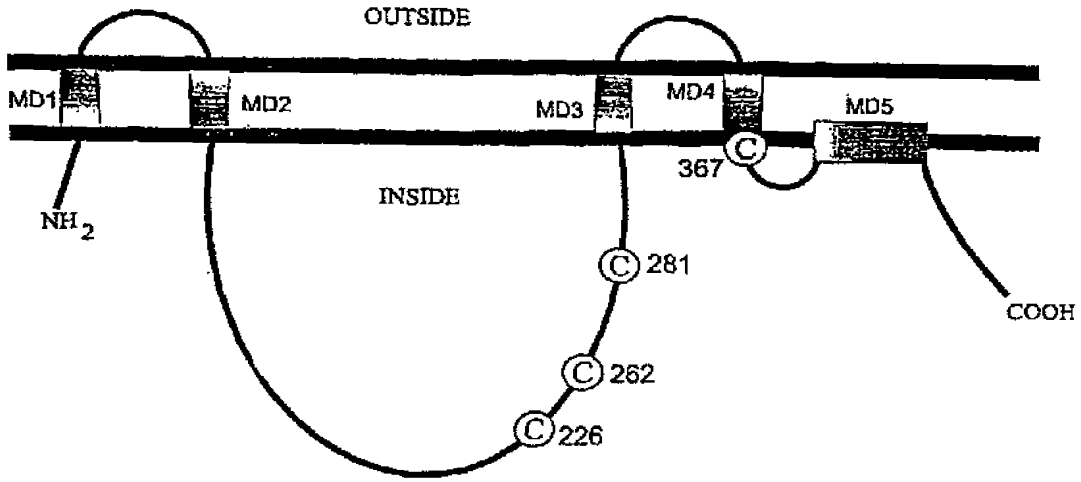
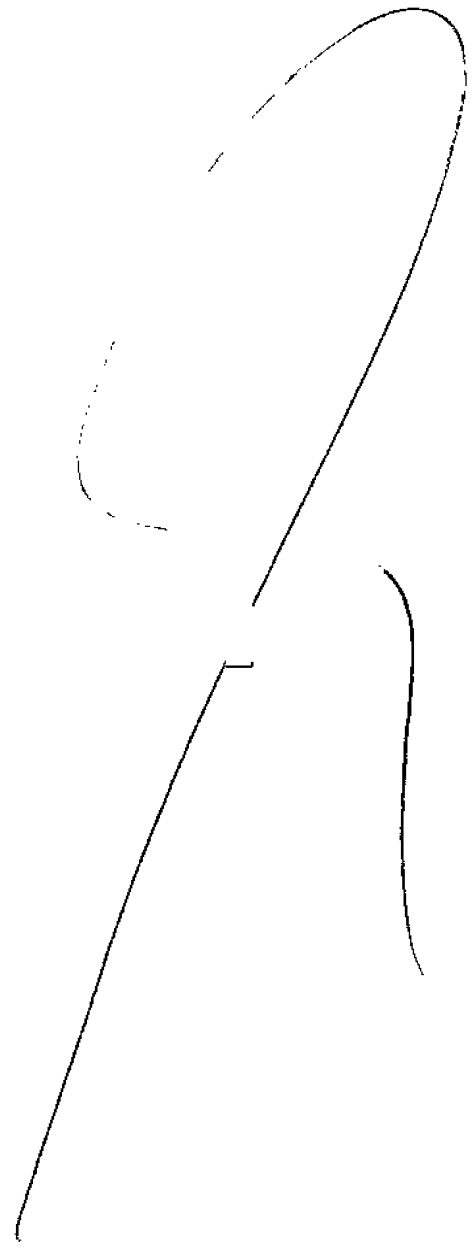
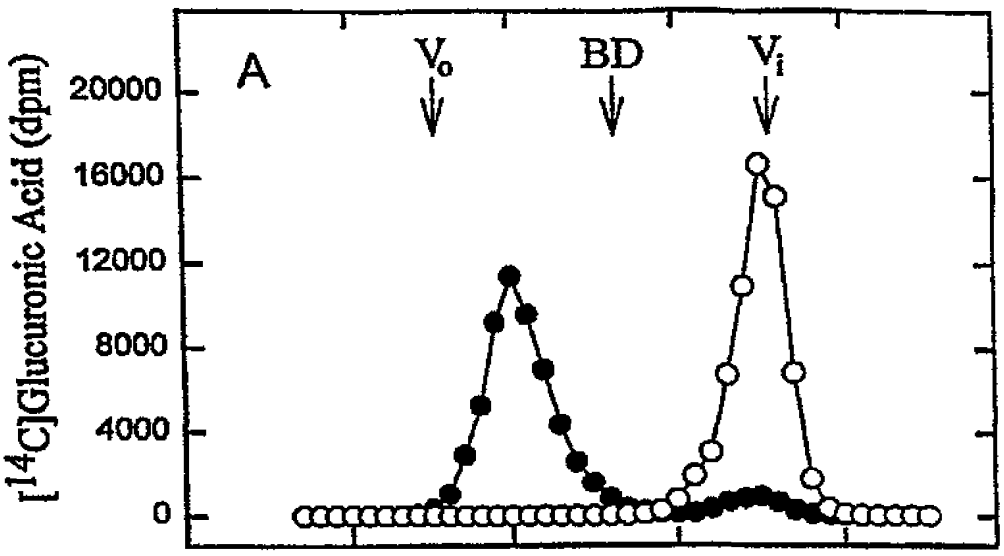


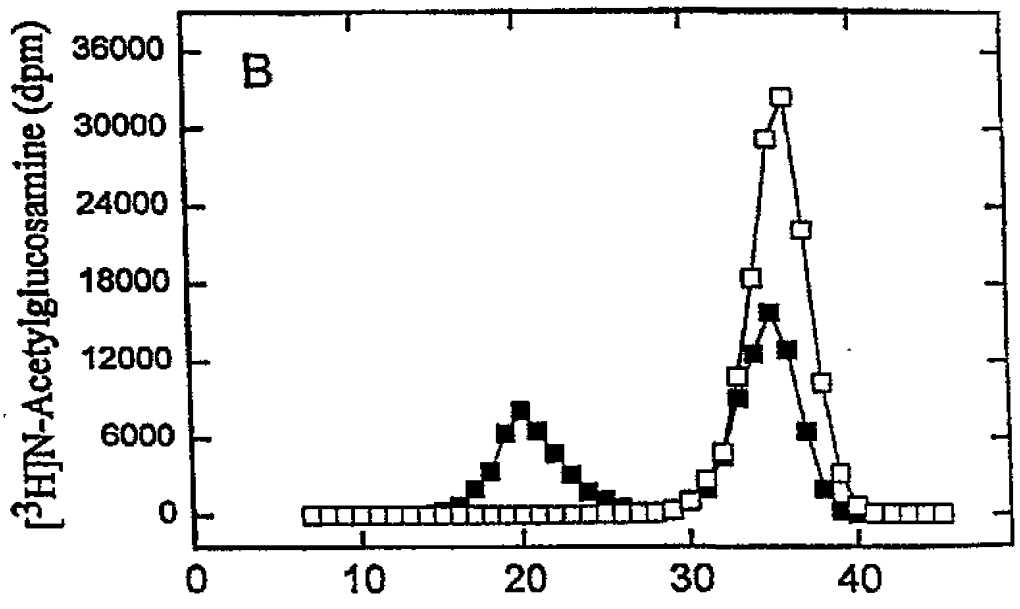
FIG. 11





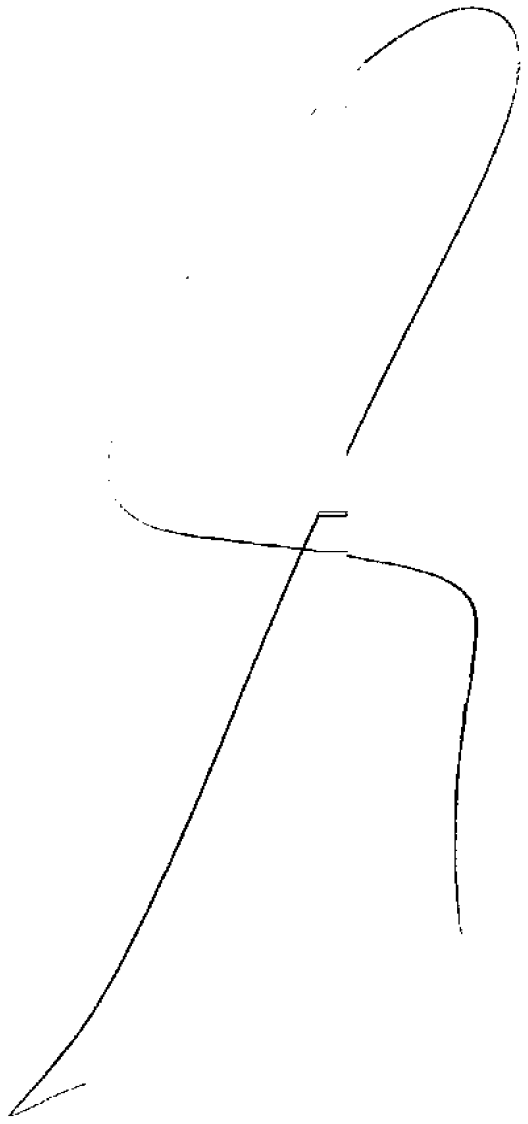
ELUTION VOLUME (ml)

Fig. 12A



ELUTION VOLUME (ml)

Fig. 12B



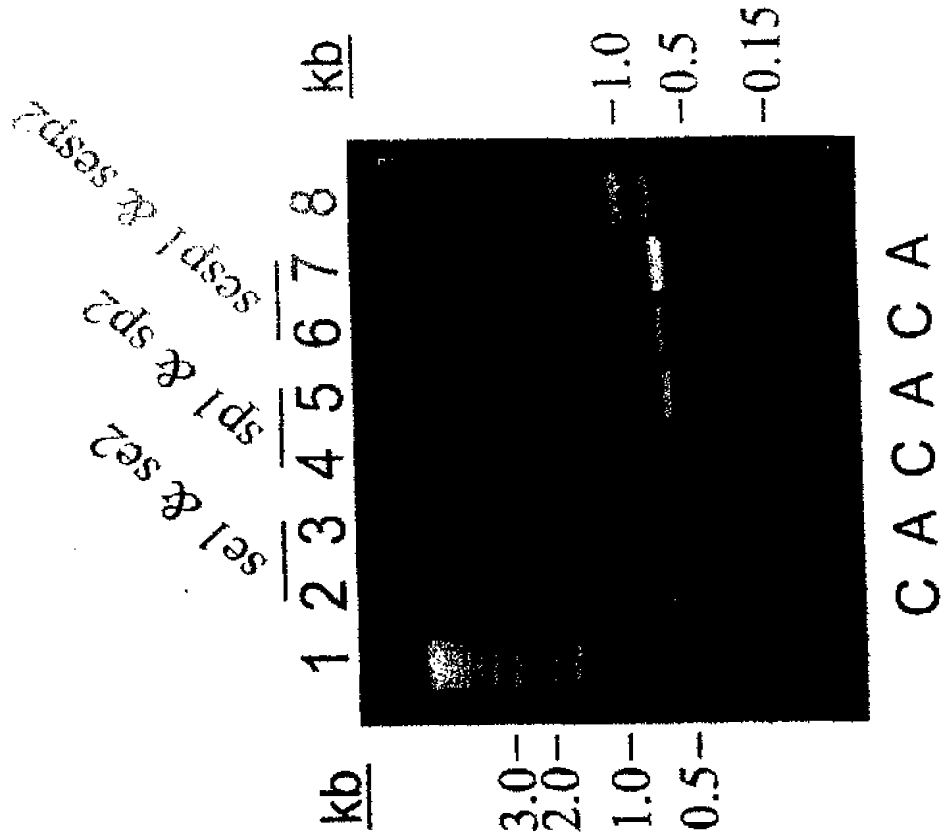
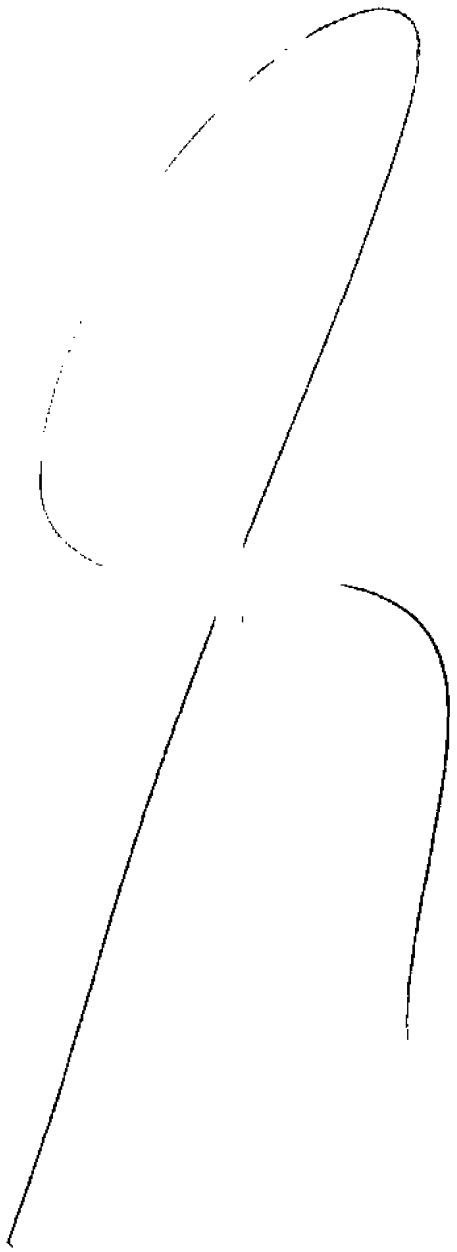
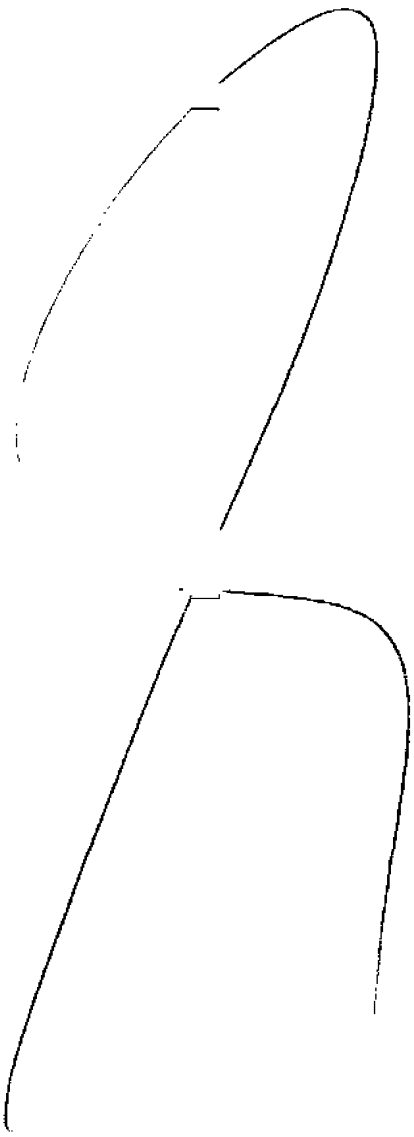

kb
1.0
0.5
0.15
1 2 3 4 5 6 7 8
sal & sp2
sp1 & sp2
sosp1 & sosp2
kb
3.0
2.0
1.0
0.5
CACACA

Fig. 13



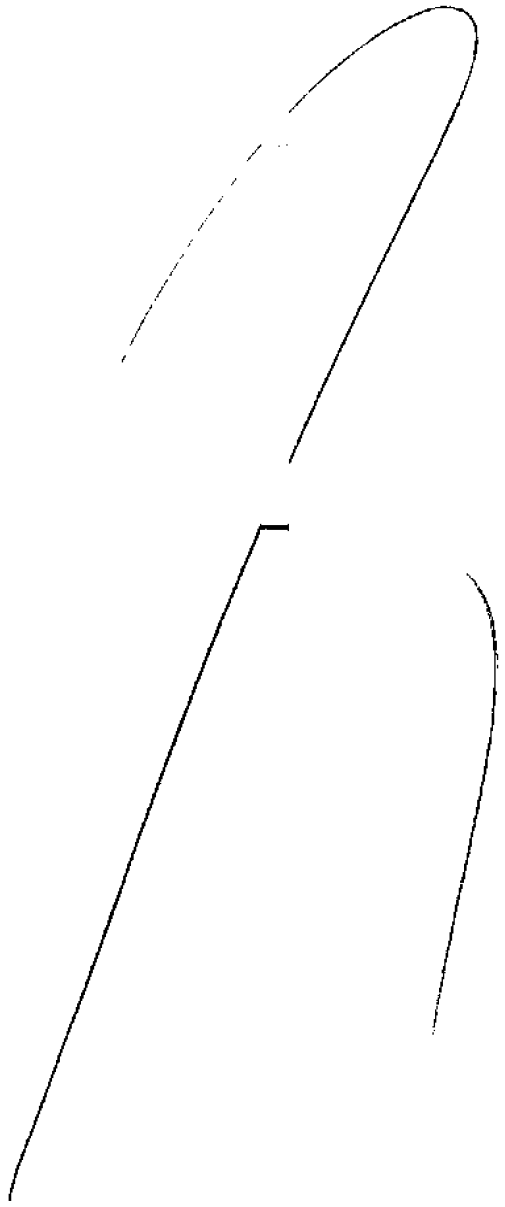
```
atgagaacattaaaaaacctcataactgttgtggcctttagtatttttgggtactgttgatttacgtcaat    72

gtttatctctttggtgctaaaggaagcttgtcaatttatggctttttgctgatagcttacctattagtcaaa   144

atgtccttatccttttttacaagccatttaagggaagggctgggcaatataaggttgcagccattattccc    216

tcttataacgaagatgctgagtcattgctagagaccttaaaaagtgttcagcagcaaacctatcccctagca   288

                                    sel →
gaaatttatgttgttgacgatggaagtGCTGATGAGACAGGTATTAAGCgcattgaagactatgtgcgtgac   360

actggtgacctatcaagcaatgtcattgttcatcggtcagagaaaaatcaaggaaagcgtcatgcacaggcc   432

                                                  sesp1 →
tgggcctttgaaagatcagacgctgatgtctttttgaccgttGACTCAGATACTTATATCTAccctgatgct   504

ttagaggagttgttaaaaacctttaatgacccaactgtttttgctgcgacgggtcaccttaatgtcagaaat   576

agacaaaccaatctcttaacacgcttgacagatattcgctatgataatgcttttggcgttgaacgagctgcc   648

caatccgttacaggtaatatccttgtttgctcaggtccgcttagcgtttacagacgcgaggtggttgttcct   720

aacatagatagatacatcaaccagaccttcctgggtattcctgtaagtattggtgatgacaggtgcttgacc   792

aactatgcaactgatttaggaaagactgtttatcaatccactgctaaatgtattacagatgttcctgacaag   864

atgtctacttacttgaagcagcaaaaccgctggaacaagtccttctttagagagtccattatttctgttaag   936

aaaatcatgaacaatccttttgtagccctatggaccatacttgaggtgtctatgtttatgatgcttgtttat  1008

                                ← se2
tctgtggtggatttctttgtagGCAATGTCAGAGAATTTGATtggctcagggttttagcctttctggtgatt  1080

atcttcattgttgccctgtgtcggaacattcattacatgcttaagcacccgctgtccttcttgttatctccg  1152

ttttatggggtgctgcatttgtttgtcctacagcccttgaaattatattctctttttactattagaaatgct  1224

          ← sesp2
gacTGGGGAACACGTAAAAAattattataa                                             1254
```

FIG. 14



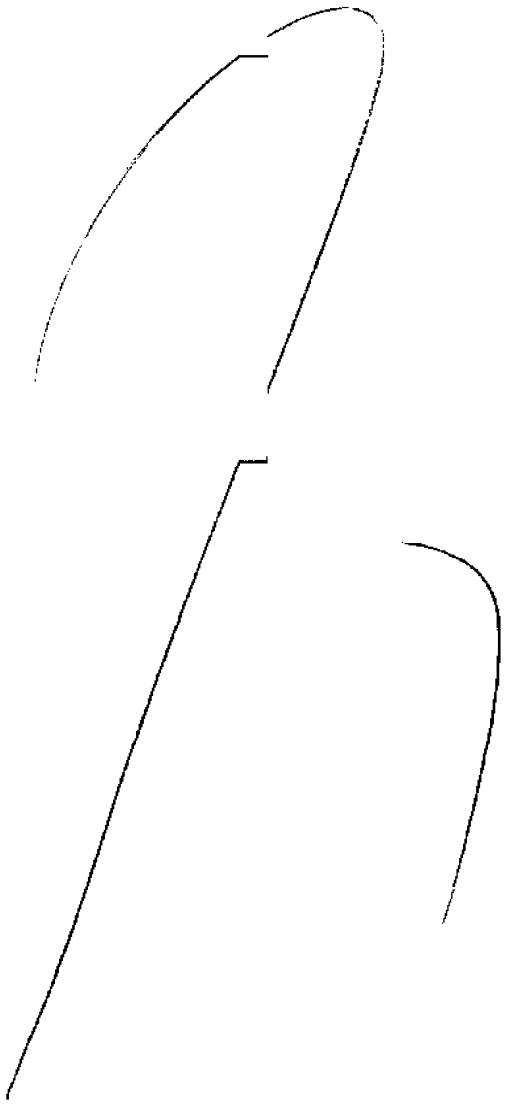
ATGAGAACATTAAAAAACCTCATAACTGTTGTGGCCTTTAGTATTTTTTGGGTACTGTTGATTTACGTCAAT	72
GTTTATCTCTTTGGTGCTAAAGGAAGCTTGTCAATTTATGGCTTTTTGCTGATAGCTTACCTATTAGTCAAA	144
ATGTCCTTATCCTTTTTTTACAAGCCATTTAAGGGAAGGGCTGGGCAATATAAGGTTGCAGCCATTATTCCC	216
TCTTATAACGAAGATGCTGAGTCATTGCTAGAGACCTTAAAAAGTGTTCAGCAGCAAACCTATCCCCTAGCA	288
GAAATTTATGTTGTTGACGATGGAAGTGCTGATGAGACAGGTATTAAGCGCATTGAAGACTATGTGCGTGAC	360
ACTGGTGACCTATCAAGCAATGTCATTGTTCATCGGTCAGAGAAAAATCAAGGAAAGCGTCATGCACAGGCC	432
TGGGCCTTTGAAAGATCAGACGCTGATGTCTTTTTGACCGTTGACTCAGATACTTATATCTACCCTGATGCT	504
TTAGAGGAGTTGTTAAAAACCTTTAATGACCCAACTGTTTTTGCTGCGACGGGTCACCTTAATGTCAGAAAT	576
AGACAAACCAATCTCTTAACACGCTTGACAGATATTCGCTATGATAATGCTTTTGGCGTTGAACGAGCTGCC	648
CAATCCGTTACAGGTAATATCCTTGTTTGCTCAGGTCCGCTTAGCGTTTACAGACGCGAGGTGGTTGTTCCT	720
AACATAGATAGATACATCAACCAGACCTTCCTGGGTATTCCTGTAAGTATTGGTGATGACAGGTGCTTGACC	792
AACTATGCAACTGATTTAGGAAAGACTGTTTATCAATCCACTGCTAAATGTATTACAGATGTTCCTGACAAG	864
ATGTCTACTTACTTGAAGCAGCAAAACCGCTGGAACAAGTCCTTCTTTAGAGAGTCCATTATTTCTGTTAAG	936
AAAATCATGAACAATCCTTTTGTAGCCCTATGGACCATACTTGAGGTGTCTATGTTTATGATGCTTGTTTAT	1008
TCTGTGGTGGATTTCTTTGTAGGCAATGTCAGAGAATTTGATTGGCTCAGGGTTTTAGCCTTTCTGGTGATT	1080
ATCTTCATTGTTGCCCTGTGTCGGAACATTCATTACATGCTTAAGCACCCGCTGTCCTTCTTGTTATCTCCG	1152
TTTTATGGGGTGCTGCATTTGTTTGTCCTACAGCCCTTGAAATTATATTCTCTTTTTACTATTAGAAATGCT	1224
GACTGGGGAACACGTAAAAAATTATTATAA	1254

SEQUENCE ID NO. 1



M	R	T	L	K	N	L	I	T	V	V	A	F	S	I	F	W	V	L	L	I	Y	V	N	24
V	Y	L	F	G	A	K	G	S	L	S	I	Y	G	F	L	L	I	A	Y	L	L	V	F	48
M	S	L	S	F	F	F	Y	K	P	F	K	G	R	A	G	Q	Y	K	V	A	A	I	I	P	72
S	Y	N	E	D	A	E	S	L	L	E	T	L	K	S	V	Q	Q	Q	T	Y	P	L	A	96
E	I	Y	V	V	D	D	G	S	A	D	E	T	G	I	K	R	I	E	D	Y	V	R	D	120
T	G	D	L	S	S	N	V	I	V	H	R	S	E	K	N	Q	G	K	R	H	A	Q	A	144
W	A	F	E	R	S	D	A	D	V	F	L	T	V	D	S	D	T	Y	I	Y	P	D	A	168
L	E	E	L	L	K	T	F	N	D	P	T	V	P	A	A	T	G	H	L	N	V	R	N	192
R	Q	T	N	L	L	T	R	L	T	D	I	R	Y	D	N	A	F	G	V	E	R	A	A	216
Q	S	V	T	G	N	I	L	V	C	S	G	P	L	S	V	Y	K	R	E	V	V	V	F	240
N	I	D	R	Y	I	N	Q	T	F	L	G	I	F	V	S	I	G	D	D	R	C	L	T	264
N	Y	A	T	D	L	G	K	T	V	I	Y	Q	S	T	A	K	C	I	T	D	V	P	D	K	288
M	S	T	Y	L	K	Q	Q	N	R	W	N	K	S	F	F	R	E	S	I	I	S	V	K	312
K	I	M	N	N	P	F	V	A	L	W	T	I	L	E	V	S	M	F	M	N	L	V	Y	336
S	V	V	D	F	F	V	G	N	V	R	E	F	D	W	L	R	V	L	A	F	L	V	I	360
I	F	I	V	A	L	I	F	N	I	F	Y	M	L	K	H	P	L	S	F	L	L	S	P	384
F	Y	G	V	L	H	L	F	V	L	Q	P	L	R	W	L	Y	S	L	F	T	I	R	N	A	408
D	W	G	T	R	K	H	L	L	*																417

SEQUENCE ID NO. 2



SEQUENCE ID NO. 3

5'-GCTGATGAG ACAG G TAT TAAGC

primer: sel (sense, nucleotides G^{316} - C^{337})

SEQUENCE ID NO. 4

5'-A T C A A A T T C T C T G A C A T T G C

primer: se2 (antisense, for sense nucleotides G^{1031} - T^{1050})

SEQUENCE ID NO. 5

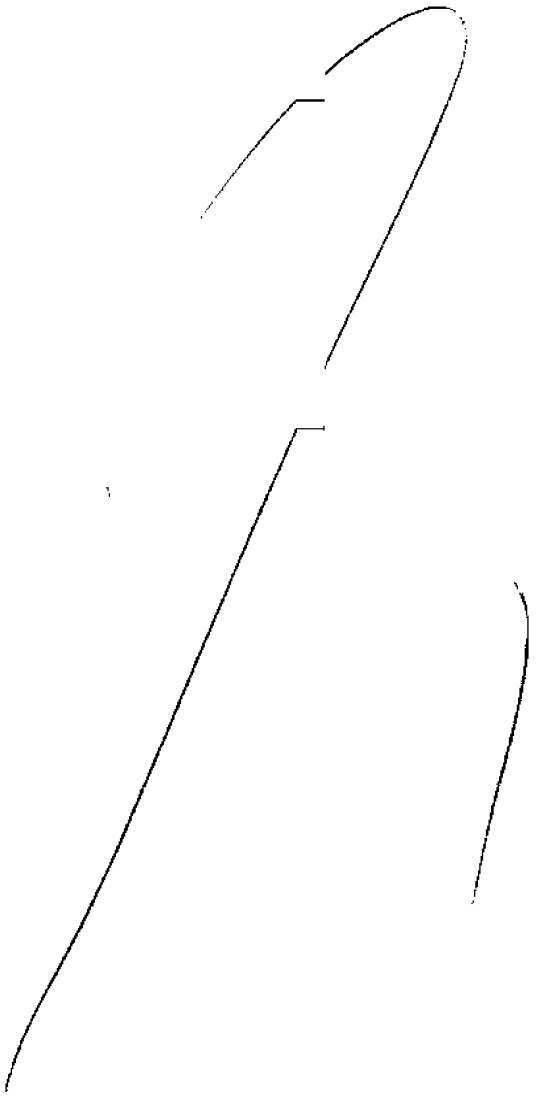
5'-G A C T C A G A T A C T T A T A T C T A

primer: sesp1 (sense, for nucleotides G^{475} - A^{494})

SEQUENCE ID NO. 6

5'-T T T T T A C G T G T T C C C C A

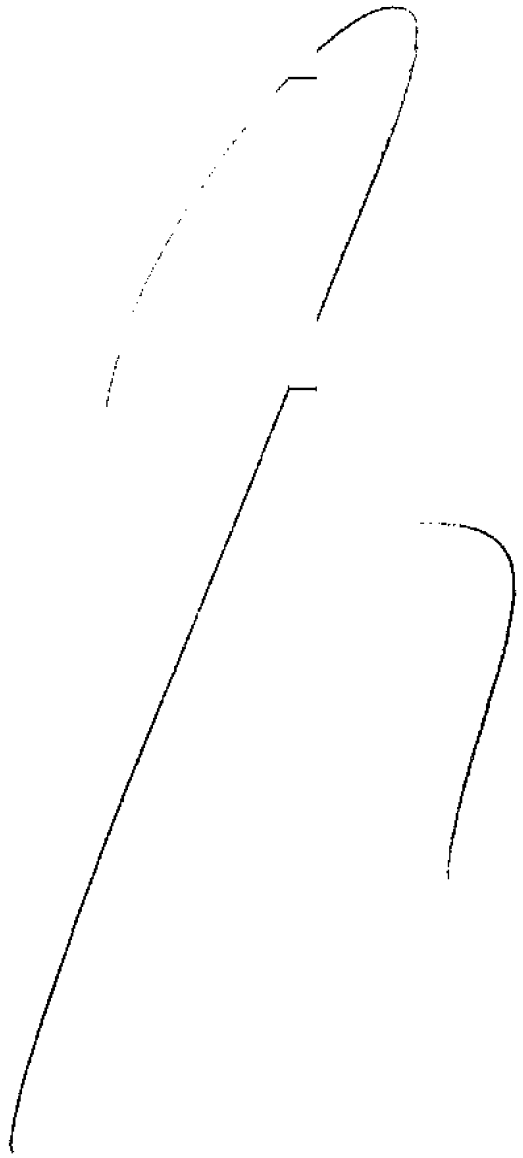
primer: sesp2 (antisense, for sense nucleotides T^{1228} – A^{1244})



Protein sequence of A98R, the PBCV-1 HA synthase

```
1     MGKNIIIMVS WYTIITSNLI AVGGASLILA PAITGYVLHW NIALSTIWGV SAYGIFVFGF
61    FLAQVLFSEL NRKRLRKWIS LRPKGWNDVR LAVIIAGYRE DPYMFQKCLE SVRDSDYGNV
121   ARLICVIDGD EDDDMRMAAV YKAIYNDNIK KPEFVLCESD DKEGERIDSD FSRDICVLQP
181   HRGKRECLYT GFQLAKMDPS VNAVVLIDSD TVLEKDAILE VVYPLACDPE IQAVAGECKI
241   WNTDTLLSLL VAWRYYSAFC VERSAQSFFR TVQCVGGPLG AYKDIIKEIK DPWISQRFLG
301   QKCTYGDDRR LTNEILMRGK KVVFTPFAVG WSDSPTNVFR YIVQQTRWSK SWCREIWYTL
361   FAAWKHGLSG IWLAFECLYQ ITYFFLVIYL FSRLAVEADP RAQTATVIVS TTVALIKCGY
421   FSFRAKDIRA FYFVLYTFVY FFCMIPARIT AMMTLWDIGW DTRGGNEKPS VGTRVALWAK
481   QYLIAYMWWA AVVGAGVYSI VHNWMFDWNS LSYRFALVGI CSYIVFIVIV LVVYFTGKIT
541   TWNFTKLQKE LIEDRVLYDA TTNAQSV
                                    567
```

SEQUENCE ID NO. 7



Nucleotide Sequence of A98R gene in the PBCV-1 Virus Genome

Start: ATG 50901 Stop: TGA 52607

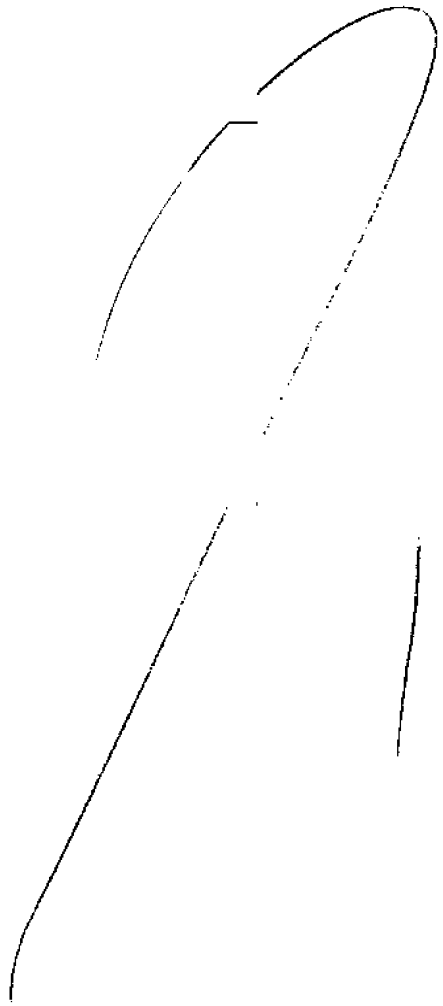
```
50881 aagacttctt gaaagttaca ATGggtaaaa atataatcat aatggtttcg tggtacacca
50941 tcataacttc aaatctaatc gcggttggag gagcctctct aatcttggct ccggcaatta
51001 ctgggtatgt tctacattgg aatattgctc tctcgacaat ctggggagta tcagcttatg
51061 gtattttcgt ttttgggttt ttccttgcac aagttttatt ttcagaactg aacaggaaac
51121 gtcttcgcaa gtggatttct ctcagaccta agggttggaa tgatgttcgt ttggctgtga
51181 tcattgctgg atatcgcgag gatccttata tgttccagaa gtgcctcgag tctgtacgtg
51241 actctgatta tggcaacgtt gcccgtctga tttgtgtgat tgacggtgat gaggacgatg
51301 atatgaggat ggctgcccgtt tacaaggcga tctacaatga taatatcaag aagcccgagt
51361 ttgttctgtg tgagtcagac gacaaggaag gtgaacgcat cgactctgat ttctctcgcg
51421 acatttgtgt cctccagcct catcgtggaa aacgggagtg tctttatact gggtttcaac
51481 ttgcaaagat ggaccccagt gtcaatgctg tcgttctgat tgacagcgat accgttctcg
51541 agaaggatgc tattctggaa gtttgtatacc cacttgcatg cgatcccgag atccaagccg
51601 ttgcaggtga gtgtaagatt tggaacacag acactctttt gagtcttctc gtcgcttggc
51661 ggtactattc tgcgttttgt gtggagagga gtgcccagtc ttttttcagg actgttcagt
51721 gcgttggggg gccactgggt gcctacaaga ttgatatcat taaggagatt aaggacccct
51781 ggatttccca gcgctttctt ggtcagaagt gtacttacgg tgacgaccgc cggctaacca
51841 acgagatctt gatgcgtggt aaaaaggttg tgttcactcc atttgctgtt ggttggtctg
51901 acagtccgac caatgtgttt cggtacatcg ttcagcagac ccgctggagt aagtcgtggt
51961 gccgcgaaat ttggtacacc ctcttcgccg cgtggaagca cggtttgtct ggaatttggc
52021 tggcctttga atgtttgtat caaattacat acttcttcct cgtgatttac ctcttttctc
52081 gcctagccgt tgaggccgac cctcgcgccc agacagccac ggtgattgtg agcaccacgg
52141 ttgcattgat taagtgtggg tatttttcat tccgagccaa ggatattcgg gcgttttact
52201 ttgtgctttа tacatttgtt tacttttttct gtatgattcc ggccaggatt actgcaatga
52261 tgacgctttg ggacattggc tgggatactc gcggtggaaa cgagaagcct tccgttggca
52321 cccgggtcgc tctgtgggca aagcaatatc tcattgcata tatgtggtgg gccgcggttg
52381 ttggcgctgg agtttacagc atcgtccata actggatgtt cgattggaat tctctttctt
52441 atcgttttgc tttggttggt atttgttctt acattgtttt tattgttatt gtgctggtgg
52501 tttatttcac cggcaaaatt acgacttgga atttcacgaa gcttcagaag gagctaatcg
52561 aggatcgcgt tctgtacgat gcaactacca atgctcagtc tgtgTGAttt ttcctgcaag
```

SEQUENCE ID NO. 8

Nucleotide and Protein Sequence of Pasteurella multocida

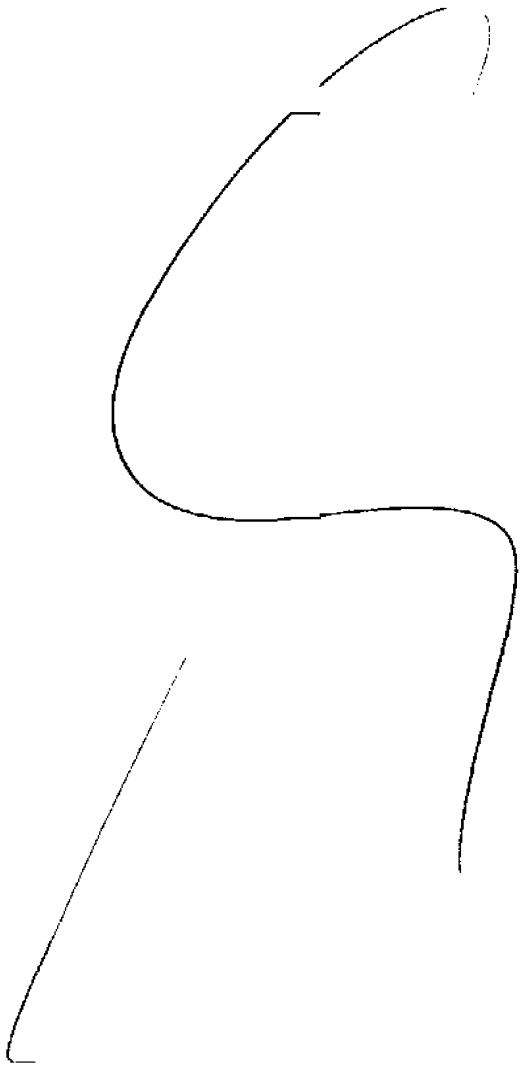
```
                                                    +1®
  1                          M  N  T  L  S  Q  A  I  K  A  Y  N  S  N  D  Y  Q
-18  ATTTTTTAAGGACAGAAAATGAATACATTATCACAAGCAATAAAAGCATATAACAGCAATGACTATCAA

 18   L  A  L  K  L  F  E  K  S  A  E  I  Y  G  R  K  I  V  E  F  Q  I  T
 52  TTAGCACTCAAATTATTTGAAAAGTCGGCGGAAATCTATGGACGGAAAATTGTTGAATTTCAAATTACC

 41   K  C  Q  E  K  L  S  A  H  P  S  V  N  S  A  H  L  S  V  N  K  E  E
121  AAATGCCAAGAAAAACTCTCAGCACATCCTTCTGTTAATTCAGCACATCTTTCTGTAAATAAAGAAGAA

 64   K  V  N  V  C  D  S  P  L  D  I  A  T  Q  L  L  L  S  N  V  K  K  L
190  AAAGTCAATGTTTGCGATAGTCCGTTAGATATTGCAACACAACTGTTACTTTCCAACGTAAAAAAATTA

 87   V  L  S  D  S  E  K  N  T  L  K  N  K  W  K  L  L  T  E  K  K  S  E
259  GTACTTTCTGACTCGGAAAAAAACACGTTAAAAAATAAATGGAAATTGCTCACTGAGAAGAAATCTGAA

110   N  A  E  V  R  A  V  A  L  V  P  K  D  F  P  K  D  L  V  L  A  P  L
328  AATGCGGAGGTAAGAGCGGTCGCCCTTGTACCAAAAGATTTTCCCAAAGATCTGGTTTTAGCGCCTTTA

133   P  D  H  V  N  D  F  T  W  Y  K  K  R  K  K  R  L  G  I  K  P  E  E
397  CCTGATCATGTTAATGATTTTACATGGTACAAAAAGCGAAAGAAAAGACTTGGCATAAAACCTGAACAT

156   Q  H  V  G  L  S  I  I  V  T  T  F  N  R  P  A  I  L  S  I  T  L  A
466  CAACATGTTGGTCTTTCTATTATCGTTACAACATTCAATCGACCAGCAATTTTATCGATTACATTAGCC

179   C  L  V  N  Q  K  T  E  Y  P  F  E  V  I  V  T  D  D  G  S  Q  E  D
535  TGTTTAGTAAACCAAAAAACACATTACCCGTTTGAAGTTATCGTGACAGATGATGGTAGTCAGGAAGAT

202   L  S  P  I  I  R  Q  Y  E  N  K  L  D  I  R  Y  V  R  Q  K  D  N  G
604  CTATCACCGATCATTCGCCAATATGAAAATAAATTGGATATTCGCTACGTCAGACAAAAAGATAACGGT

225   F  Q  A  S  A  A  R  N  M  G  L  R  L  A  K  Y  D  F  I  G  L  L  D
673  TTTCAAGCCAGTGCCGCTCGGAATATGGGATTACGCTTAGCAAAATATGACTTTATTGGCTTACTCGAC

248   C  D  M  A  P  N  P  L  W  V  R  S  Y  V  A  E  L  L  E  D  D  D  L
742  TGTGATATGGCGCCAAATCCATTATGGGTTCATTCTTATGTTGCAGAGCTATTAGAAGATGATGATTTA

271   T  I  I  G  P  R  K  Y  I  D  T  Q  H  I  D  P  K  D  F  L  N  N  A
811  ACAATCATTGGTCCAAGAAAATACATCGATACACAACATATTGACCCAAAAGACTTCTTAAATAACGCG

294   S  L  L  E  S  L  P  E  V  K  T  N  N  S  V  A  A  K  G  E  G  T  V
880  AGTTTGCTTGAATCATTACCAGAAGTGAAAACCAATAATAGTGTTGCCGCAAAAGGGGAAGGAACAGTT

317   S  L  D  W  R  L  E  Q  F  E  K  T  E  N  L  R  L  S  D  S  P  F  R
949  TCTCTGGATTGGCGCTTAGAACAATTCGAAAAAACAGAAAATCTCCGCTTATCCGATTCGCCTTTCCGT

340   F  F  A  A  G  N  V  A  F  A  K  K  W  L  N  K  S  G  F  F  D  E  E
1018 TTTTTTGCGGCGGGTAATGTTGCTTTCGCTAAAAAATGGCTAAATAAATCCGGTTTCTTTGATGAGGAA
```

SEQUENCE ID NO.19
(Page 1 of 3)



```
 363   F  N  H  W  G  G  E  D  V  F  P  G  Y  R  L  F  R  Y  G  S  F  F  K
1087   TTTAATCACTGGGGTGGAGAAGATGTGGAATTTGGATATCGCTTATTCCGTTACGGTAGTTTCTTTAAA

 386   T  I  D  G  I  M  A  Y  H  Q  E  P  P  G  K  E  N  E  T  D  R  E  A
1156   ACTATTGATGGCATTATGGCCTACCATCAAGAGCCACCAGGTAAAGAAAATGAAACCGATCGTGAAGCG

 409   G  K  N  I  T  L  D  I  M  R  E  K  V  P  Y  I  Y  R  K  L  L  P  I
1225   GGAAAAAATATTACGCTCGATATTATGAGAGAAAAGGTCCCTTATATCTATAGAAAACTTTTACCAATA

 432   E  D  S  H  I  N  R  V  P  L  V  S  I  Y  I  P  A  Y  N  C  A  N  Y
1294   GAAGATTCGCATATCAATAGAGTACCTTTAGTTTCAATTTATATCCCAGCTTATAACTGTGCAAACTAT

 455   I  Q  R  C  V  D  S  A  L  N  Q  T  V  V  D  L  E  V  C  I  C  N  D
1363   ATTCAACGTTGCGTAGATAGTGCACTGAATCAGACTGTTGTTGATCTCGAGGTTTGTATTTGTAACGAT

 478   G  S  T  D  N  T  L  E  V  I  N  K  L  Y  G  N  N  P  R  V  R  I  M
1432   GGTTCAACAGATAATACCTTAGAAGTGATCAATAAGCTTTATGGTAATAATCCTAGGGTACGCATCATG

 501   S  K  P  N  G  G  I  A  S  A  S  N  A  A  V  S  F  A  K  G  Y  Y  I
1501   TCTAAACCAAATGGCGGAATAGCCTCAGCATCAAATGCAGCCGTTTCTTTTGCTAAAGGTTATTACATT

 524   G  Q  L  D  S  D  D  Y  L  E  P  D  A  V  E  L  C  L  K  E  F  L  K
1570   GGGCAGTTAGATTCAGATGATTATCTTGAGCCTGATGCAGTTGAACTGTGTTTAAAAGAATTTTTAAAA

 547   D  K  T  L  A  C  V  Y  T  T  N  R  N  V  N  P  D  G  S  L  I  A  N
1639   GATAAAACGCTAGCTTGTGTTTATACCACTAATAGAACGTCAATCCGGATGGTAGCTTAATCGCTAAT

 570   G  Y  N  W  P  E  F  S  R  E  K  L  T  T  A  M  I  A  H  H  F  R  M
1708   GGTTACAATTGGCCAGAATTTTCACGAGAAAAACTCACAACGGCTATGATTGCTCACCACTTTAGAATG

 593   F  T  I  R  A  W  H  L  T  D  G  F  N  E  K  I  E  N  A  V  D  Y  D
1777   TTCACGATTAGAGCTTGGCATTTAACTGATGGATTCAATGAAAAAATTGAAAATGCCGTAGACTATGAC

 616   M  F  L  K  L  S  E  V  G  K  F  K  H  L  N  K  I  C  Y  N  R  V  L
1846   ATGTTCCTCAAACTCAGTGAAGTTGGAAAATTTAAACATCTTAATAAAATCTGCTATAACCGTGTATTA

 639   H  G  D  N  T  S  I  K  K  L  G  I  Q  K  K  N  H  F  V  V  V  N  Q
1915   CATGGTGATAACACATCAATTAAGAAACTTGGCATTCAAAAGAAAAACCATTTTGTTGTAGTCAATCAG

 662   S  L  N  R  Q  G  I  T  Y  Y  N  Y  D  E  F  D  D  L  D  E  S  R  K
1984   TCATTAAATAGACAAGGCATAACTTATTATAATTATGACGAATTTGATGATTTAGATGAAAGTAGAAAG

 685   Y  I  F  N  K  T  A  E  Y  Q  E  E  I  D  I  L  K  D  I  K  I  I  Q
2053   TATATTTTCAATAAAACCGCTGAATATCAAGAAGAGATTGATATCTTAAAAGATATTAAAATCATCCAG

 708   N  K  D  A  K  I  A  V  S  I  F  Y  P  N  T  L  N  G  L  V  K  K  L
2122   AATAAAGATGCCAAAATCGCAGTCAGTATTTTTTATCCCAATACATTAAACGGCTTAGTGAAAAAACTA

 731   N  N  I  I  E  Y  N  K  N  I  F  V  I  V  L  H  V  D  K  N  H  L  T
2191   AACAATATTATTGAATATAATAAAAATATATTCGTTATTGTTCTACATGTTGATAAGAATCATCTTACA

 754   P  D  I  K  K  E  I  L  A  F  Y  H  K  H  Q  V  N  I  L  L  N  N  D
2260   CCAGATATCAAAAAAGAAATACTAGCCTTCTATCATAAACATCAAGTGAATATTTTACTAAATAATGAT
```

SEQUENCE ID NO.19
(Page 2 of 3)



```
 777   I  S  Y  Y  T  S  N  R  L  I  K  T  E  A  H  L  S  N  I  N  K  L  S
2329   ATCTCATATTACACGAGTAATAGATTAATAAAAACTGAGGCGCATTTAAGTAATATTAATAAATTAAGT

 800   Q  L  N  L  N  C  E  Y  I  I  F  D  N  H  D  S  L  F  V  K  N  D  S
2398   CAGTTAAATCTAAATTGTGAATACATCATTTTTGATAATCATGACAGCCTATTCGTTAAAAATGACAGC

 823   Y  A  Y  M  K  K  Y  D  V  G  M  N  F  S  A  L  T  H  D  W  I  E  K
2467   TATGCTTATATGAAAAAATATGATGTCGGCATGAATTTCTCAGCATTAACACATGATTGGATCGAGAAA

 846   I  N  A  H  P  P  F  K  K  L  I  K  T  Y  F  N  D  N  D  L  K  S  M
2536   ATCAATGCGCATCCACCATTTAAAAAGCTCATTAAAACTTATTTTAATGACAATGACTTAAAAAGTATG

 869   N  V  K  G  A  S  Q  G  M  F  M  T  Y  A  L  A  H  E  L  L  T  I  I
2605   AATGTGAAAGGGGCATCACAAGGTATGTTTATGACGTATGCGCTAGCGCATGAGCTTCTGACGATTATT

 892   K  E  V  I  T  S  C  Q  S  I  D  S  V  P  E  Y  N  T  E  D  I  W  F
2674   AAAGAAGTCATCACATCTTGCCAGTCAATTGATAGTGTGCCAGAATATAACACTGAGGATATTTGGTTC

 915   Q  F  A  L  L  I  L  E  K  K  T  G  H  V  F  N  K  T  S  T  L  T  Y
2743   CAATTTGCACTTTTAATCTTAGAAAAGAAAACCGGCCATGTATTTAATAAAACATCGACCCTGACTTAT

 938   M  P  W  E  R  K  L  Q  W  T  N  E  Q  I  E  S  A  K  R  G  E  N  I
2812   ATGCCTTGGGAACGAAAATTACAATGGACAAATGAACAAATTGAAAGTGCAAAAAGAGGAGAAAATATA

 961   P  V  N  K  F  I  I  N  S  I  T  L  *
2881   CCTGTTAACAAGTTCATTATTAATAGTATAACTCTATAA
```

SEQUENCE ID NO.19
(Page 3 of 3)